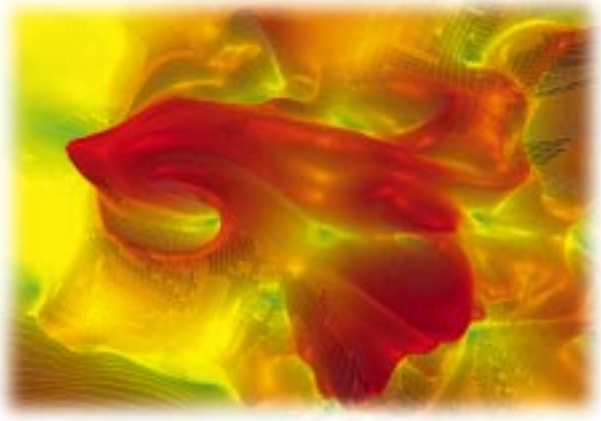
BITS

computing & communications news

Members of the Advanced Computing Lab's Systems Team install the latest Extreme Linux cluster called Rockhopper. See inside front cover for more information.

Los Alamos
NATIONAL LABORATORY

Feb/March 2000



Results from the Applied Physics Division's RAGE radiation-hydrodynamics code. The image shows density variations in the mixing zone generated by the passage of a shock through an irregular interface between materials (Richtmyer-Meshkov instability). Both height and color indicate density. For more information, contact Richard Holmes, holmes@lanl.gov , (505) 667-3598; Robert Weaver, rpw@lanl.gov , (505) 667-4756; and Michael Gittings, gittings@lanl.gov ,(505) 667-1453, all from the Thermonuclear Applications Group (X-2).

About the front cover: Rockhopper, the latest Extreme Linux cluster at LANL's Advanced Computing Laboratory, plays the dual role of an open computing platform and a testbed for computer science research. The initial Rockhopper cluster consists of four scalable units, for a total of 128 dual PIII 500-MHz D-nodes. Each D-node has 1 GByte of RAM, a hot-swappable 9-GByte disk (with room for additional hot-swappable disks), and is connected to both the 100baseT and Myrinet fabrics. The front cover photos show from top to bottom: John Patchett uncrating a rack of the new system; Patchett, Susan Coghlan, and Ron Minnich node testing and installing Myrinet cards; Coghlan and Dean Prichard working at the back of the cluster; and a close-up view of two of Rockhopper's racks. For more information see the article in this issue or contact Coghlan at smc@lanl.gov.

Produced by the Computing, Information,
and Communications (CIC) Division
Design: Donald Montoya and Julie Medina (CIC-1)
Managing Editor: Denise Sessions, denise@lanl.gov (CIC-1)
Illustration: Dave Delano, Art/Illustration
Printing: Imaging Services Group (CIC-9)
HTML: David Van Etten (CIC-1)

Attention Subscribers: There are two ways you can update your subscription data. E-mail bitsupdate@lanl.gov, or fill out the form inside this issue. Internal LANL subscribers please note that until we have a mechanism in place to verify our subscription database with the LANL Employee Information System we will rely on you to update your mailstop. Thanks for keeping us updated on subscription data.

Table of Contents

Blanca/Tecolote Base Class Redesign

Hitting the Wall: When the Past Rises Up to Haunt You

by John H. Hall, Technical Staff Member, Applied Physics Division, X-10: Code Group B

Suppose you've been developing a computer science framework for slightly more than three years and you've had numerous successes in demonstrating an unprecedented ease of use, portability, parallel efficiency, and simplicity in the creation of new models. You're feeling pretty good about your accomplishments and honestly believe you have discovered the best thing since sliced bread.

Then you are asked to add a new model, proposed by people who look at the world from a different perspective —a different galaxy, really. As they describe their requirements, a sinking feeling in the pit of your stomach begins to eat at you. What these folks want to do is fairly simple; in fact, you are able to write a rapid prototype of a stand-alone code in a matter of days. However, for you to implement this simple model in your framework will require weeks of redesign because you had already decided (with much worry and discussion at the time) to implement a basic assumption slightly differently than that required for this new task.

To add to your humiliation, suppose you had just written an article entitled "Achieving Revolution Through Evolution," in which you described how a lot of planning will help you overcome these little obstacles by using a series of small changes to eventually get you back on track. Now what???

This humbling scenario has just played itself out in my own life: not once, but three times in the past few months. Each time I have seen that decisions we had made to reduce the amount of information our end users needed for input had limited our flexibility in a critical area. We had hit a wall and had to find some way to get over or around it.

A Week in the Library…

My thesis adviser used to repeat incessantly that a week in the library is worth six months in the laboratory. At the time, this statement did little more than annoy me, but I have since realized that most of us—faced with daily pressures—decide to make cuts in personal-growth areas. We skip an issue of our favorite technical magazine, we don't leave our offices for a visiting speaker's presentation or a lunchtime talk with our colleagues, and (most important!) we come to regard the library as a bastion of solitude for only those with nothing better to do. The primary reasons for failing to learn about others' past experiences are the same reasons touted for why we can't reuse software: I can do it myself faster and with less effort than would be required to learn somebody else's implementation. ***This is an untruth so profound in both its magnitude and implications for our society that it rivals war as a waste of resources***.

Blanca/Tecolote development has been dedicated to determining how to write infrastructure code precisely—once for all time. This approach meant the code had to be of high quality as well as efficient and versatile. Now we were being told we hadn't quite made the grade on the versatile part. Youch, that stings!!!

Under normal circumstances I would have immediately tried to fix the problems in whatever manner came to mind. But I had just been reading Tom DeMarco's *The Deadline,* which deals with project management. The mythical project he describes has some teams that are too large to work together efficiently under the imposed structure (this plot follows ideas set forth in Frederick Brook's *Mythical Man Month,* which claims any team greater than ten in size is too large).

Unfortunately, DeMarco's teams could not be subdivided because Management-on-High would interpret that move to mean that staff had less than full commitment to achieving the preordained goals of the project. This dilemma put middle-to-low-level managers between a rock and a hard place if they felt that the true measure

of their management efforts was determined by the productivity of the team. The characters knew that by dividing the large team into smaller ones, they could achieve greater productivity and yet they might be accused of not giving the corporate goals their full support. Actually, as long as they delivered on time, no one cared because everybody would try to share the credit. The problem arose when—through perception or fact—the project was declared to be in trouble.

When a problem is under duress, the statement "If only (s)he had utilized the entire team, they could have delivered" crops up with irritating frequency. Although politically useful to the uninvolved, this statement is babble with no basis in fact. People caught using this phrase should be forced to wash their mouths out with soap and sent upstairs to bed with no supper. (I actually have much stronger ideas on suitable punishments, but I figured I couldn't get them by the editors, so why bother enumerating them in elegant detail.)

The entire Blanca/Tecolote team comprises more than 30 people—some full time and some part time on the project—from the fields of computer science, physics, and engineering. Obviously, such a team is difficult to assemble frequently, and it is too diverse to achieve consensus quickly.

My solution to the thorny problem of how to divide the team was to hold a series of *nonmandatory* design team meetings, for which topics were announced in advance and an open invitation was extended. Our team members—all being rational individuals who try to minimize the number of meetings they attend—self-selected into subteams, only coming to those meetings for which the topics were of urgent and direct interest to them. No formal teams were ever announced and yet we had achieved the subdivision necessary to be able to arrive at appropriate decisions. *Cautionary Note: the only way this technique can work is if everyone is so busy they feel they must skip any meeting that is not absolutely necessary. If your team members do not feel this way, they are not busy enough.*

The original problem that spawned these design meetings was a lack of flexibility in parts of our framework. In the previous design, flexibility was primarily limited by the paucity of decision points we used to control the flow of the program. The task of increasing the number of potential decision points in object-oriented programming boils down to redesigning the base classes in the inheritance hierarchy. After the base classes are redesigned, the new public interface must be propagated to the derived classes. Thus in adding flexibility, we established more locations in which to add or change code, and each of the derived classes needs to be repackaged to take advantage of these new opportunities. For a project like Blanca/Tecolote, which has a large code base, this is a big job. Furthermore, the customers already using our codes would not appreciate our going away for a few months because we would not be available to fix any problems they might encounter. Nevertheless, the time had come to make a midcourse correction, there was no simple way to put it off and we had to act.

The Base Class Redesign

The original base classes in Blanca/Tecolote had three opportunities to be called by the framework:

- during the constructor,
- in the *initialize* method after the persistents were loaded, and
- in the *eval* method once per cycle.

The significant work usually happened in the *eval* method. Often we had to wait to initialize some variables until the *eval* method was called because the variables depended on other objects' existence, which could not be guaranteed during the *initialize* method. These delayed initializations were then needlessly tested during each subsequent pass through the method.

Following my thesis adviser's library credo (research the existing material *first*), I examined what other component architectures had done in the past. One of the best and most long-lived was the Adobe PhotoShop™ "plug-in" architecture, which had created a subservient industry of third-party developers that supplied a variety of add-ons for the most popular graphics program of all time. Entire sophisticated applications were written using Adobe's plug-in architecture. This was the kind of generality Blanca/Tecolote needed. Of course, Adobe's early efforts required a little tuning to handle the new possibilities presented by a modern object-oriented language (they had used an early Apple Computer language known as Object Pascal), but John and Tom Knoll of Adobe essentially had the calling sequence right.

In the latest redesign we have now expanded the calling sequence to include:

- the constructor;
- persistent loading;
- the *initialize* method;
- a *PreRun* method;
- a loop over a *Start* method, a *Continue* method, and a *Finish* method; and
- a *PostRun* method. (See Fig. 1.)

In addition, we can also subcycle the *Continue* method.

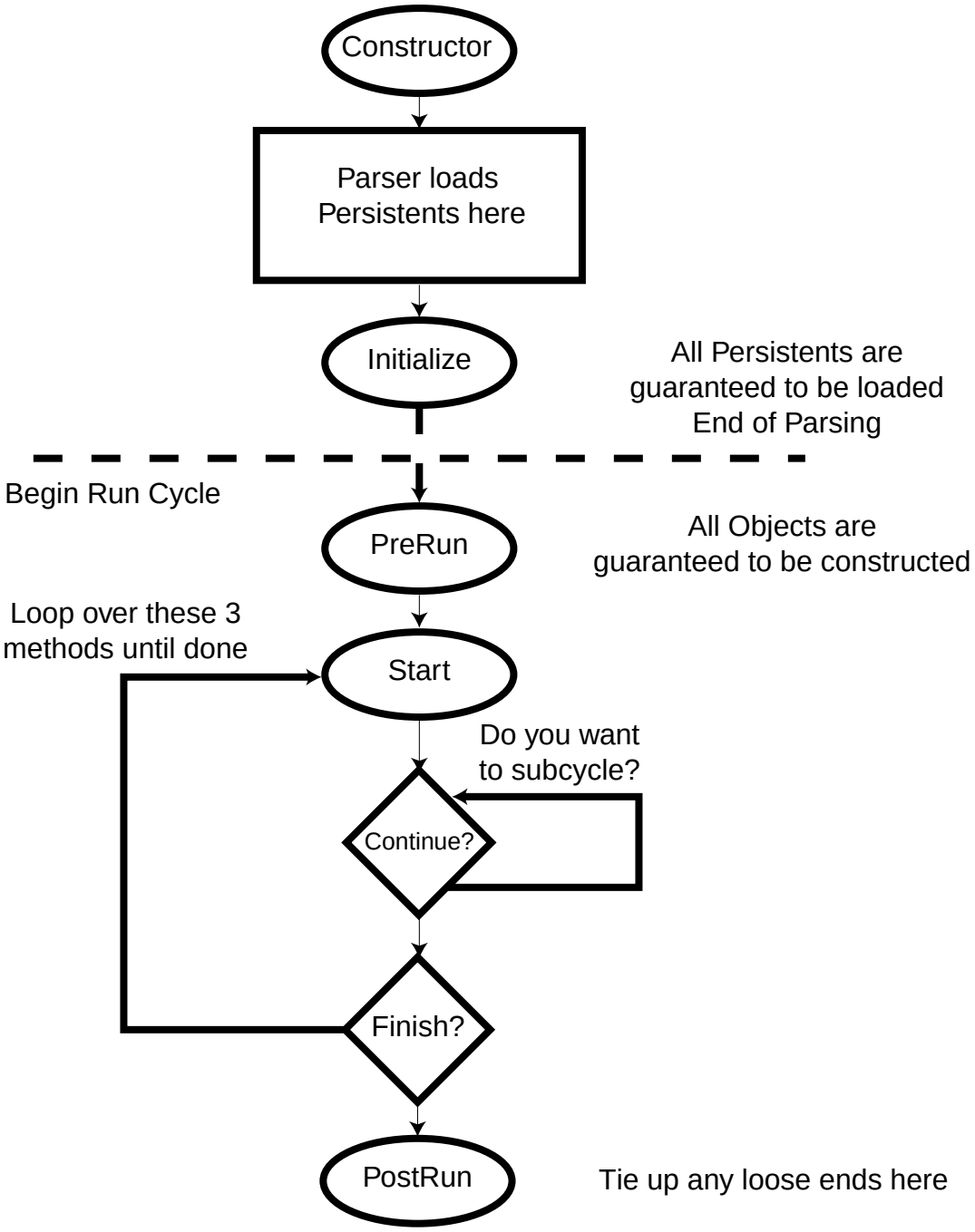


Fig. 1. Program Flow after the Base Class Redesign.

By breaking the original *eval* method into these five separate functions, we actually factorize the initialization of objects so that we no longer must conditionally test for them. Any unused functions are simply removed by the optimizing linker now available in modern compilers, so these extra abstractions do not reduce efficiency.

Another real boon is that the existing code does not need to be rewritten—just repackaged. Although this process will involve a lot of scripting, cutting, and pasting, it is by no means as complex as writing the original code. The flexibility achieved in this design actually simplifies the coding. If, for instance, you want to initialize a static counter inside the inner loop, just set it to zero in the *Start* method and then use it in the *Continue* and *Finish* methods. During each pass through the loop, the counter is reset correctly by the *Start* method. Under normal circumstances, initializing static counters involves a complicated set of *if* tests that require careful consideration of the "which came first, the chicken or the egg" issue. In this new scheme, it is obvious where to initialize and the process is simple.

Hiding the complex computer science that underlies the Blanca/Tecolote framework requires constant vigilance. During this repackaging, we are constantly asking the question, "How can we simplify the user interface?" For example, we saw one area for significant simplification in the multi-material (MM) and single-material (SM) model class hierarchies. When we introduced the concept of *Recursive* materials, the differences between MM models and SM models were reduced to a theoretical exercise that unnecessarily complicated the framework. We removed this now artificial distinction and simplified to a single model base class that serves both purposes. Once again, our efforts fall in the repackaging rather than complete-rewrite category.

Summary

From the beginning of the Blanca/Tecolote Project, we've endorsed the premise that a framework is not really a framework until at least three distinct applications have been constructed with it. We are now reaping the benefits of insight gained from the many disparate applications being built using our framework. Unfortunately, a side effect of such a development-on-the-fly approach is that decisions—seemingly harmless enough in the beginning—sometimes come back to haunt you. Figuring out how to move forward without inconveniencing an installed user base is often very difficult. However, taking the time for a periodic review of your history and goals, followed by a period of tuneups to correct problems, is extremely beneficial. Perhaps it is appropriate that this first Blanca/Tecolote article of the new year (and for zero-based counting systems, the new millennium) should employ a historical perspective.

Acknowledgment

I would like to acknowledge the contributions of Jean Marshall, Mark Zander, and the other Blanca/Tecolote Team members who have vigorously entered into the Base Class Redesign debate. Their careful consideration of the issues has helped us plot our course into the future.

Editor's note: BITS is featuring a series of articles on Blanca/Tecolote. These articles have been published in previous issues, which can be accessed online at http://www.lanl.gov/orgs/cic/cic6/bits/archive.html .

- *Tecolote: An Object-Oriented Framework for Hydrodynamics Physics*, August 1999.
- *Using Tecolote Components to Extend Object-Oriented Programming*, September 1999.
- *Frameworks are Models, Too!*, October 1999.
- *Dimensionless Coding Techniques in Tecolote: Using a Single Source-Code Base for Multidimension Programs*, November 1999.
- *Achieving Revolution Through Evolution: Breaking Giant Leaps into Small Steps,* January 2000.

Rockhopper: Extreme Linux at the ACL

by Susan Coghlan, Technical Staff Member, Advanced Computing Laboratory

With the advent of Linux and the availability of low cost commodity PC hardware, clusters of PCs running Linux have become popular as an economical way to provide large-scale computation for a fraction of the cost of traditional "big machines." In the past year, many have jumped on the Linux bandwagon, touting it as the solution to all computing problems. But, as one would expect with a young operating system, Linux is not perfect, especially in a high-performance computing (HPC) environment, such as the Advanced Computing Laboratory (ACL). For the past four years, researchers at the ACL have been investigating the use of commodity hardware and Linux for HPC, building successively larger "Extreme Linux" clusters and working with the community to develop Linux as an HPC operating system. Extreme Linux clusters can be considered a super class of Beowulf clusters.

Beowulf clusters are large, fast machines built from cheap, commodity pieces. Scientists at the Goddard Space Flight Center brought them into the mainstream around 1994. Typically these clusters are simply piles of PCs, built by stacking many Intel Pentium desktop PCs together on shelves, installing identical free operating systems across all of them, and wiring them together with 100baseT and inexpensive switches. The significant difference between a Beowulf cluster and a bunch of networked PCs sitting on individual user desks is the software, which allows computers in the cluster to efficiently work together toward a common goal. Using this software, based on the industry standard Message Passing Interface (MPI), a user can run a single large job across all the nodes in the cluster, gaining access to increased overall CPUs, memory, and disk resources.

A System Rivaling the Supercomputers of the Past

Beowulf clusters have received a great deal of media attention. Part of their appeal is their low cost when compared to the price of a similar capability machine from a traditional supercomputer vendor. But Beowulf clusters are also spreading rapidly because they function extremely well for a specific class of applications. These applications, such as N-body simulations, use embarrassingly parallel algorithms, which have infrequent communications, and, as such, are well suited to running on many fast nodes with a moderately high-latency/slow interconnect. However, the high latency and relative slowness of TCP/IP over 100baseT, and the general lack of scientific computing tools in Linux, has prevented these clusters from being heavily used in communication-intensive HPC applications. This lack of tools and the poor interconnect provide the primary motivation for what we call "Extreme Linux" clusters.

With Extreme Linux clusters, we attempt to combine the best of both Beowulf clusters and traditional supercomputers. By using low-cost commodity hardware, fast low-latency interconnects, and by providing an environment rich in tools and supporting applications for the scientist, we hope to build machines that work as well, or better, than standard supercomputers.

With Extreme Linux clusters, we attempt to combine the best of both Beowulf clusters and traditional supercomputers. By using low-cost commodity hardware, fast low-latency interconnects, and by providing an environment rich in tools and supporting applications for the scientist, we hope to build machines that work as well, or better, than standard supercomputers. However, the difference between a large single-image (one copy of the operating system running across all processors), symmetric multiprocessor (SMP) supercomputer and a cluster of many small SMP or uniprocessor (UP) machines is much more than just the difference in cost and operating system. For efficient use of these machines, the programming model must be different from the one used for the typical supercomputing platform. An MPI code will run on both architectures, but if you care about performance you may not want to do message passing within an SMP node, instead you might use MPI between nodes and lightweight threads inside each node. Another obvious difference between the clusters and the traditional supercomputer is in the area of administration and management of the system image. Having so many individual system images to keep up-to-date and consistent across the cluster increases the complexity of managing the system by an order of magnitude. Compounding the problem is the use of an operating system that is still fairly young and was optimized for small PCs.

Open Source Software Is a Big Advantage

Although Linux is an immature operating system, it has one feature that gives it a big advantage over proprietary operating systems used for supercomputers—its kernel and base programs are all "open source" software. The term open source means that the source code is available in its entirety for debugging, patching, and extending by anyone, without the need for them to pay license fees, sign nondisclosure agreements (NDAs), or involve lawyers. Having the source code available to many independent researchers encourages them to pore over the code, fixing bugs, testing out new algorithms, optimizing for their applications, etc.

This ability to modify the Linux kernel and system software is well suited to HPC because HPC applications tend to stress operating systems in ways that standard computer applications do not. Having an open source operating system allows researchers and administrators working on an HPC system to investigate key sections of the code, and to extend, tune, and optimize the operating system for their unique applications and algorithms. Because the basic operating system is open source, and because there is a very large community of developers working on it, Linux has improved so rapidly in the past few years that it is becoming a serious contender in the supercomputing world. In fact, at least one supercomputer vendor has embraced Linux as a replacement for the proprietary operating system they currently use on their supercomputers. In addition, Linux has a small, well designed, stable, and mature kernel that is less than 300 K lines of code (without device drivers). This provides a level of robustness and flexibility that greatly increases its suitability for supercomputing.

Linux as a Production HPC Operating System

However, while Linux has proven to work well for Internet, technical, and server applications, and while it offers many advantages over other, more traditional HPC operating systems, it is not quite ready for production HPC deployment. Initial tests on our Extreme Linux clusters have highlighted weaknesses in Linux that must be resolved before Linux-based clusters can be used as production HPC platforms.

Extreme Linux Brings Together Collaborative Research

In 1998, the ACL began the Extreme Linux effort, and brought together researchers in the field to identify common shortcomings and strategic directions for high-performance Linux. We also began working closely with representatives from Myricom, Compaq, Red Hat, Portland Group, IBM, and SGI on their high-performance Linux strategy. Together with other laboratories, we identified the components missing for terascale SMP Linux clusters and have been attempting to focus our efforts on the key issues that must be resolved before these clusters are capable of high-performance computation.

The New Rockhopper Cluster

The latest ACL Extreme Linux cluster (Rockhopper) arrived at the beginning of this year. It was purchased with funding from a diverse set of sources within LANL, including a significant investment from Deputy Director Bill Press's office. This cluster plays the dual role of an open computing platform and a testbed for computer science research.

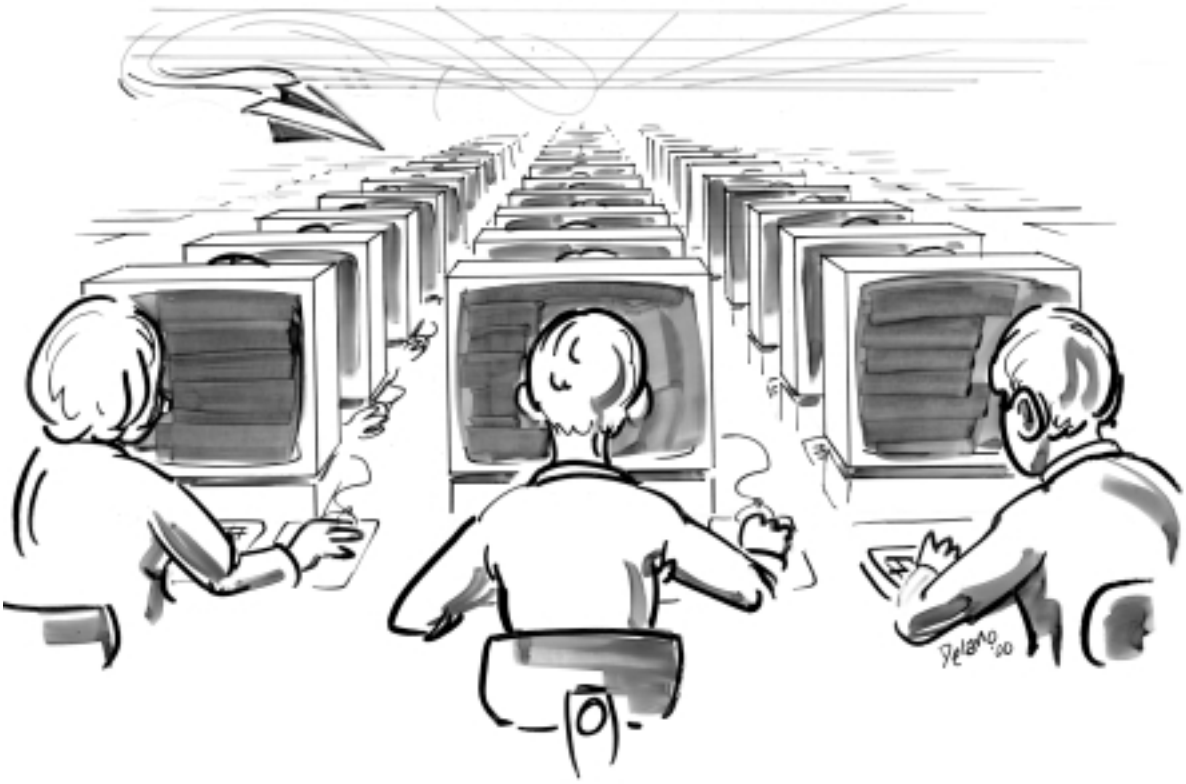
Scalability was the fundamental design concept for Rockhopper. We needed to have the ability to start with a small cluster and grow to a multiteraflop cluster in a short time period. Our experiences with previous clusters had made it clear that this would not be a trivial task. (See miniarticle following this article called "The History of ACL's Research with Extreme Linux Clusters.")

The Sandia National Laboratories' Cplant research project, so far the most advanced study of large-scale PC clusters, has made important contributions in area of cluster scalability. Their research has shown that a hierarchical architecture is required for clusters much larger than 100 nodes. Building on Sandia's work and working in collaboration with Argonne National Laboratory, we developed the concept of what we call a "scalable unit" or SU. (For a complete description of an SU, see the Scalable Unit sidebar.)

The initial Rockhopper cluster consists of four SUs, for a total of 128 dual PIII 500-MHz D-nodes. Each D-node has 1 GB of RAM, a hot-swappable 9-GB disk (with room for additional hot-swappable disks), and is connected to both the 100baseT and Myrinet fabrics. All 128 D-nodes of the Rockhopper cluster are currently online and running benchmark, test, and "friendly user" codes.

The computer science research on the Rockhopper cluster includes collaborations with other national laboratories in administration, management tool development, and in scalable cluster design. The ACL is working with Argonne, National Energy Research Scientific Computing Center (NERSC), and VA Linux to develop methods for managing system images, installing hundreds of nodes quickly, updating large groups of nodes easily, and booting with alternate system images on a per job basis. The System Imager package from VA Linux provides a

(*Continued on page 9*)



Scalable Unit Sidebar

A "scalable unit," or SU, is a self-contained unit of rack-mounted SMP D-nodes (dynamic nodes), connected to other SUs and external networks with one or more high-speed network interfaces. The D-nodes will be the compute nodes and will be dedicated to running user applications. They are "dynamic" because, for a terascale cluster, the status of the machine is constantly in flux, and we expect that the nodes will become unavailable for periods of time for a variety of reasons. Software must be able to dynamically schedule jobs from a pool of operational D-nodes. Each SU contains a Power, Hardware status, and Interconnect monitoring node (PHI node). The PHI node is the control/server node for the SU and is used to monitor all aspects of the SU, including the state of the D-Nodes within the unit, and any faults or routing problems within the high-speed interconnect fabric. PHI nodes provide remote access to the "consoles" of the D-Nodes, and also provide system images for installing and dynamically updating the D-Nodes as needed. User programs do not run on the PHI nodes, they are strictly reserved for control and server use only. The PHI nodes are able to reset and power cycle each D-Node within their SU, and are also used for job scheduling, managing the nodes, etc. As the cluster grows, nodes for managing the PHI nodes can be added, hopefully preserving scalability for up to tens of thousands of processors. An important design element of the SUs includes the ability to quickly, and easily, replace whole nodes without disruption of jobs running on other nodes.

Each SU consists of 32 D-nodes and a fast 100baseT switch so that nodes within an SU are directly connected to each other. The 100baseT switch is then connected to the other SUs, and external networks, through dual gigabit Ethernet uplinks. Our research indicated that having two interconnect fabrics worked well, using TCP/IP over 100baseT for the control traffic that is unable to effectively utilize low-latency hardware, and reserving the more expensive high-speed low-latency hardware for message passing traffic. The cost-effectiveness of the new 64+64 (128 port) Myrinet Clos switches combined with the fact that Myricom now supports a Linux driver (GM) and provides a Linux GM MPI implementation, made Myrinet the obvious choice for our high-performance interconnect.

mechanism to install, modify, and update a system image on the fly. This package, combined with research by Ron Minnich, James Hendricks, and Dale Webster of the ACL, will provide a quick and efficient method for switching nodes from one operating system/kernel to another. These researchers have Linux booting Linux (without going through the BIOS) and have installed the Open BIOS on PCs in the ACL computer test lab. They are now working on modifying the Open BIOS to hold a small Linux kernel and to allow boots from that kernel. In effect, these modifications replace the PC BIOS with Linux.

The Rockhopper cluster has multiple metaservers and compile/front-end machines. Each of these machines has a large RAID array attached. We are currently running early release versions of the ReiserFS and Ext3 journaled file systems on the RAID arrays. In addition to providing access to the Rockhopper cluster, these machines are used for our research in high-performance file systems. At the ACL, we are collaborating with Stelias Computing to provide extensions to InterMezzo, a file system with useful cluster characteristics. This work involves implementing the automatic migration of user files to cached file systems as needed and building a cached root file system. As part of the work with Stelias Computing, we are also designing a robust cluster file system and integrating Private Name Space semantics. InterMezzo will eventually replace NFS for most purposes on our cluster.

As part of the work on administration and management of large clusters, researchers at the ACL are developing tools capable of monitoring large numbers of nodes efficiently. Two such tools, Supermon and Superview, written by Ron Minnich, Karen Reid, and Matt Sottile, can monitor hundreds of nodes at rates of up to 100 Hz. Beta versions of these tools have demonstrated the ability to monitor at 10 Hz without significant impact on running applications. The tools monitor hardware performance counters and collect a wide range of kernel information (disk blocks, memory, interrupts, etc). The kernel was modified so all the parameters can be grabbed without going through /proc (a special area used to store information about the system, running processes, etc). The Power, Hardware status, and Interconnect monitoring nodes (PHI nodes) then collect and process the information from the D-Nodes. The PHI node is the control/server node for the SU and is used to monitor all aspects of the SU, including the state of the D-Nodes within the unit, and any faults or routing problems within the high-speed interconnect fabric. (See the SU sidebar for more information about the PHI nodes.) By gathering the collective data from the PHI nodes, a user's program can be monitored. The design of these tools allows queries or even "triggers" to be written. The intent is that the PHI nodes would be able to monitor the health of the individual D-nodes and attempt to repair and/or reboot nodes as needed. Because of this ability to track the status of D-nodes, PHI nodes would be able to remove unresponsive D-nodes from the scheduler's pool of available compute nodes. The Rockhopper cluster will be used to test and develop these tools further for scalability and robustness.

Challenges Remain

Using the Rockhopper cluster, we are continuing our research in zero-copy messaging (See miniarticle on The History of Extreme Linux at the ACL). Experiences on the Blue Mountain machine demonstrated that communication between SMPs could be extremely complex. We expect terascale clusters of SMPs to have multiple smart network adapters and the capability to directly transfer messages between SMPs without making time-consuming copies of the data. Smart adapters can allow messages to be delivered without disrupting the CPU cache or stealing cycles from the CPU.

However, current strategies for allocating buffer space for messages, pinning down memory, reserving sections of the pinned memory for transfer, and identifying data corruption all begin to fail as the number of processors participating in the message passing reach the thousands. Probably the largest unresolved problems are buffer management, data corruption, and retransmission. Most of the methods available for zero-copy messaging use a technique called "sender managed buffering," in which the sender of the message assumes it manages a small portion of the remote RAM reserved for zero-copy messaging.

Unfortunately, this technique does not scale. For example, a 2000-node cluster would require that each node keep track of 1999 remote buffer areas. Even with a modest 4-MB message area per send-receive pair, the scheme would require network adapters to reserve 8 GB of RAM for messaging—possibly more memory than used by the actual application! The Sandia Cplant research work has provided a scalable messaging layer, but at the price of simplifying the problem—zero copy messaging is abandoned. In addition to these difficulties, there is the problem of how to deal with the occasional corrupted message. Detecting a corrupt message is not difficult, but devising a scalable scheme for resending the correct data remains a research project. Peter Beckman and the ACL run-time team are using the Rockhopper cluster to investigate several novel techniques that they believe could be used to solve these problems with zero-copy messaging between smart network adapters.

Providing a Production Quality Computing Environment

Also, continuing efforts from previous ACL Extreme Linux clusters, we hope to further develop a production quality scientific computing environment. Although we now have many tools, such as the Totalview parallel debugger, ported to Linux, many are still lacking in user support. We will continue to provide and test commercial parallel development support tools as they become available for Linux. But we are also looking into some of the more prosaic system functions that often make production computation difficult for the end user.

These include tasks like quickly and efficiently distributing a user's binaries, libraries, and data files to and from the specific compute nodes allocated to their job. This is particularly difficult in HPC environments where the input and output files can be terabytes in size. Previous research on our Linux clusters has shown that using standard NFS (Network File System) between a fileserver and a cluster does not scale well. Even with the high-speed gigabit interconnection fabric, there is a serial bottleneck at a single interface between the file server and the cluster. We hope to take advantage of the multicast feature of TCP/IP over standard Ethernet and the advanced file system research described earlier in this article to provide a solution to this problem. With IP subnet broadcasting, files can be distributed with two transfers, one from the file servers to the PHI nodes, and one from the PHI nodes to their D-nodes. With just these two transfers, the user environment can be sent to any subset of the cluster participating in the computation.

Further research is planned in integrating this efficient file transfer into the job queuing software so that a user's environment can be staged to all their scheduled nodes before their job begins. Although the low-latency fabric cannot use this file transfer mechanism (owing to optimization for point-to-point communication and lack of knowledge of the IP or Ethernet protocols), we are investigating the possibilities for using Myrinet to attach disks to the cluster with network-attached storage technologies.

Hardware Handicaps

Of course, even with all the improvements that this research brings to Linux, we are still handicapped by the Intel hardware typically used for these clusters. Factors like the 32-bit architecture, low memory bandwidth, poor bus architectures, problematic BIOSes, no standard netboot support, and all that goes with maintaining backward compatibility still cause problems. However, it is not clear that using the admittedly better architecture of systems like the Compaq Alphas is a big enough win to make up for the higher cost of these workstation quality machines. The ACL has set up a small lab of Alpha systems running Linux. We plan on using this small cluster to investigate the cost effectiveness of building clusters of these workstations. Whether HPC applications can take advantage of the hardware superiority of this class of machines is an open question. In addition, Alpha Linux tends to lag behind Intel Linux with regard to porting tools, packages, and drivers. It remains to be seen whether Compaq's announced support of Linux on the Alpha hardware will accelerate work in this area.

Gaps Remain

Given all the success of Extreme Linux clusters, there are still some areas where they have a long way to go to catch up with other HPC systems. For example, in the areas of reliability, availability, and serviceability (RAS), Linux clusters running on Intel hardware still lag behind the offerings of many supercomputing vendors. Efforts such as the High Availability Linux project, recent advances in hot-swappable hardware components, and system imaging software that can largely automate replacement of faulty compute nodes help close this gap. But, other areas need more work, such as the lack of kernel dump analysis tools. Having an open source kernel makes these tools somewhat less critical than they are on proprietary operating systems, but they are still extremely valuable for debugging the sort of problems one often encounters with HPC systems.

Future Rockhopper Plans

The Rockhopper cluster will be used in research areas of job scheduling, partitioning, and accounting. The ACL systems team is transferring the work done on the Nirvana cluster to the Rockhopper cluster. Future plans for the Rockhopper cluster include expansion to 1000+ nodes, investigations into how to seamlessly integrate different generations of hardware, and research into incorporating different architectures such as the Alphas, IA64, etc. Of course, since almost all the ACL research has depended heavily on the ability to modify, extend, and optimize the Linux kernel and operating system, the ACL will continue to be involved in the open source movement.

As Extreme Linux clusters continue to spread and research into high-performance computing on these clusters continues to advance, we expect that within a few years, Linux will be the HPC operating system of choice. Because the ACL's mission is to investigate state-of-the-art high-performance computing, to provide insight into future computing environments, and to perform computer science research in areas of high-performance computation, we will be involved in Extreme Linux cluster research for the foreseeable future.

The History of ACL's Research with Extreme Linux Clusters

For the past four years, the ACL has been investigating the use of Linux clusters as a high-performance computing (HPC) platform. We started with the Peak cluster—8 dual Pentium II (PII) processor PCs, interconnected with both Myrinet and 100baseT. This small cluster was used to evaluate whether commodity hardware and Linux could provide an environment for HPC. Much of the research on this cluster was focused on low-latency interconnects. At that time, there weren't many options available for low-latency network hardware because most vendors were not interested in providing drivers for a low market-share operating system such as Linux. Fortunately, one product, Myrinet from Myricom, Inc.—a low-latency, high-speed interconnect (1.2 gigabit per second bi-directional switched), allowed users to program the interface directly. This ability to program the network interface allowed us to study the use of operating system bypass methods to avoid high overhead protocols such as TCP/IP.

At the time, Myricom didn't supply OS bypass drivers and utilities for Linux. The ACL partnered with the Real World Computing Project, a Japanese research organization, to build OS bypass libraries and utilities (named SCORE), and a Myrinet driver (PM) for our Linux cluster. Initial results were very promising, the measured latency was surprisingly low and the performance of standard benchmarks was excellent across the cluster. While it was obvious that Linux was a long way from a production HPC environment, we felt that there was enough potential to warrant further investigation.

Little Blue Penguin—The Second ACL Linux Cluster

As we have seen with Accelerated Strategic Computing Initiative (ASCI) Blue Mountain and ASCI Blue Pacific, terascale clusters quickly magnify the role of any unscalable component. Often, pieces of code that work well on a small cluster, or in "normal" use, fail miserably as the cluster scales to thousands of nodes. For example, with the next ACL Linux cluster, the Little Blue Penguin cluster (LBP), we ran into a simple problem with a commonly used piece of software. The LBP cluster consists of 64 dual processor PII PCs. When we attempted to run a standard MPI program using mpirun across the cluster, we ran up against the line-length limitations of the Unix shell. This is because mpirun uses the Unix remote shell (rsh) command to spawn processes onto remote nodes. The MPI program sends a command to the remote nodes that names each of the nodes participating in the job and the port that the program will use to listen for connections. The command line built for our 64 nodes was longer than the maximum length allowed by our version of tcsh and the code failed to run. Of course, it was trivial to modify the shell to allow longer command lines, but the real issue was not the size of the argument buffers, but that mpirun was not designed with scalability in mind. As clusters reach 5,000, 10,000, or more nodes, it will no longer be trivial to fix these types of problems. This was not the only example where we ran into difficulties caused by scaling issues.

Standard network services, such as TCP/IP, NFS, and NIS, are not as robust or as optimized in Linux as they are in other HPC operating systems. For example, the TCP/IP network would hang when MPI jobs flooded network sockets with lots of small packets. After some investigation, we were able to patch the kernel and eliminate the problem. NIS would fail intermittently and user accounts would no longer be valid on the nodes. Once again, we were able to easily resolve the problem by removing NIS and using other mechanisms to provide the same service. We also experienced problems where jobs would die when multiple nodes attempted to perform parallel reads of files located on the NFS server. Setting up local scratch areas and copying the data files to each node before running the jobs took care of this problem, albeit at a significant cost to the user. Other similar problems showed up on a regular basis. However, we were able to debug and fix each of these problems quickly because Linux is open source, whereas with a proprietary operating system we would have had to wait for the vendor to provide a patch, if they ever did. Nevertheless, it was clear that we were just squeaking underneath the scalability bar.

Exploring the SMP Frontier

The LBP cluster was also used to research the applicability of building Extreme Linux clusters using SMP nodes. Most of the Linux clusters up to this time were clusters of single processor nodes, and it was unclear whether the use of SMPs could be cost effective, or even whether Linux was capable of running correctly on SMPs nodes. At the time the LBP cluster was built, SMP locking in the Linux kernel was a disaster, causing nodes booted with SMP kernels to

crash regularly. Fortunately, the next version of the kernel included a redesign of the locking code, and within a few months the SMP kernels became reasonably stable. We still experienced occasional SMP-related problems but were able to run and benchmark codes. The benchmark results indicated that SMP nodes could be cost effective for some scientific applications. Of course, price-performance is largely dependent on the style and ratio of communication and computation performed by an application. Some communication-dependent programs will benefit from the lower latency and higher internal bisection bandwidth of an SMP, while some compute-bound applications will work fine on very cheap nodes with moderate latencies between each CPU. However, it was clear from the benchmark results, that there was still work to be done in the programming environment if we wished to effectively use both intranode and internode communication paths.

Providing the Tools

Along with the research in basic services and SMP issues, we were also investigating limitations in the programming environment. Because the motivation for installing Linux clusters was typically the result of a small budget, it was common to find clusters with no commercial software installed. However, we realized that money would have to be invested in software if we wanted an environment that would not just allow scientists to run their applications, but would actually support and help them develop and run their codes. To that end we purchased the KAI compilers, and worked closely with both KAI and the EGCS developers (providing funding to the EGCS project). In addition, we actively encouraged vendors of other commercial utilities to port their packages to Linux. As a result, Etnus ported Totalview to Linux (a production release is now available). At that time, not all needs could be met with the commercial software available for Linux. The messaging layer, middleware, and other software components necessary for HPC were missing. Helping to round out the scientific computing environment, we ported many of the ACL software packages to Linux, including SMARTS, POOMA, PETE, Tulip, Tau, PAWS, and Siloon.

The Next Step

By early summer 1999, with the improvements to Linux and increasing support from the Linux community for HPC, it was clear that Linux had the potential to work quite well as a supercomputer operating system. However, it was also apparent that there were still weaknesses, particularly in the areas of scalability, management, applications, and I/O. Some of these were limitations of the Intel hardware and some were limitations of the system software. In either case, the LBP cluster was obviously not large enough to push these limits, and a larger cluster was needed if we wished to investigate these issues.

By late summer of 1999, we started looking into the possibility of installing a large cluster. The central design choice was the ability to scale up as additional funds became available. In September of 1999, the first scalable units of the Rockhopper cluster were ordered.

New Electronic Resources Available at the Research Library

Keeping Current with Alerts

by Kathy Varjabedian, Databases Team Member, CIC-14, Research Library

Alerts are an automatic current awareness service. Alerts can be a good way to keep up with a topic of interest to you without having to take the time effort to go out and search—the information is delivered to your mailbox. Typically you register for the service and then create a search strategy or personal profile which will be matched automatically against each update of the database. An e-mail notification is sent to you on a regular basis.

You may use alerts to:

- keep current with the latest research on a topic,
- track the latest research by a particular author or institution,
- receive the table of contents of your favorite journal (from library databases), and
- see who has cited a paper or author (SciSearch® or Social SciSearch® databases only).

Research Library database alerts, offering notification of recently published literature, are

BIOSIS® at LANL

SciSearch® at LANL

Engineering Index® at LANL

Social SciSearch® at LANL

INSPEC® at LANL

How to Set Up an Alert in a Library Database

1. Go to the database of your choice (you may use the links in the list above).
2. Try out a desired search and review the results. For a topical search, a good practice is to do the search with a date limit of one year and see if the number of search results is a reasonable amount to get in weekly or biweekly alerts. Broaden or narrow the search as needed.
3. Click on alerts in the database toolbar.
4. If this is your first alert, register as a new user; otherwise, login with your user name and password.
5. Select “New Alert” (“New General Alert” or “New Cited Alert” for SciSearch®/Social SciSearch®).
6. Give your alert a name, select method of receipt, and enter search criteria as tested.
7. Submit the alert.

You receive e-mail each week, either the citations themselves or a URL for logging in to your account to view the results.

There are also other free alert services, available freely on the Web or to LANL researchers:

Fatbrain.com	Online seller of technology and science books. “Keep Me Posted” alert service offers daily, weekly, or monthly e-mail notification of new titles in one or more of 700 subject categories.
FEDIX	FEDIX Opportunity Alerts provide notification of new funding opportunities from eight participating government agencies, using a personal profile set up using their Grants Keyword Thesaurus®.
GartnerGroup Interactive	Information technology and computer science research and analysis. Be notified when GartnerGroup research is published that matches your alert, here called a Profile. Requires user name/password for the LANL account (available here), and personal registration. Select “My Home Page” to get to Profiles.
Melvyl	Alerts are available in the Melvyl catalog, MEDLINE, Computer Articles, Magazine and Journal Articles, and Newspaper Articles databases to LANL staff. Contact the Library Service Desk (7-5809 or library@lanl.gov) for a password and instructions.
Northern Light	General World Wide Web search engine. Narrow search requests are the most useful as a result of the high volume of information being searched.
SciCentral	Science and engineering news and resources. Weekly notification from any of 120+ subject areas.
Scout Report for Science and Engineering	A biweekly publication offering a selection of new and newly discovered Internet resources of interest to researchers and educators in the physical and life sciences.

Corporate Portals—Your View Into the Virtual Workplace World

by Denise Sessions, BITS Managing Editor, and Katherine Norskog, aha! Team Member, CIC-1 Communication Arts & Services

Imagine that getting the information you need is only two mouse-clicks away. That's one of the goals of the Information Management (IM) Focus Group in the Computing, Information, and Communications (CIC) Division at Los Alamos.

"The CIC Information Management Focus Group is all about helping our customers see information services as a coherent unity rather than a mere collection of service offerings," said Bob Newell, CIC deputy division director for information services. "One of our first efforts, the administrative resources Web page, organizes useful information in a single, easy to use manner."(See http://int.lanl.gov/enterprise/ .)

"But we don't want to stop there. A subteam of the IM Focus Group has been tasked with identifying the next steps for the administrative resources page that will make it grow into a full-fledged 'enterprise portal'."

The goal is to develop information management tools that turn data into information by not only organizing and categorizing data, but giving it meaning and getting it to people who need it to do their jobs. Portals are a powerful tool to organize vast quantities of data to improve information delivery and access.

Turning Data into Information

The LANL Web currently hosts over a million pages. It's as though someone emptied the contents of many file cabinets onto the floor. The information is there—somewhere—but you have to wade through many documents to find it.

Searching

Before portals came on the scene, the standard technique for providing access to a body of content was to put a full-text search engine on top of it. But a full-text search can overwhelm a user with the large number of "documents" returned, making it difficult to retrieve the most relevant information in the top results. David Curle, of Upgrade magazine (the magazine for the Software & Information Industry Association) says, "Even the search engine and document management companies are starting to include categorization tools in their products."

aha!-Directory for a Corporate Portal

aha!, an effort to add a categorization tool (and browse capabilities) to our search tool is available when you click on the magnifying glass icon or the word "search" off the Laboratory homepages. A small team of two librarians, a designer, and a programmer joined to create this Yahoo®-like tool so that the outside world could more easily locate good information quickly on the Laboratory's science and technology research.

To begin aha!, the team developed a first draft of top-level categories and queried users for feedback. They surveyed 247 individuals representing all Laboratory divisions to arrive at 14 top-level categories of LANL science and technology Web pages.

How Does It Work?

Imagine you want information on nonlinear stress analysis at the Laboratory. You can browse the category tree (taxonomy) under Analysis and Testing or type in the words nonlinear stress analysis to initiate a search. Currently, you get two technical papers from the Research Library. You also get the results of the full-text search, which is initiated at the same time. It yields 23 results including press releases, Newsbulletin stories, and agendas.

aha! includes categorized Web pages, technical reports, patents, databases, audiovisual files, and images. The team added a LANL-only version when the yellow/green network split was implemented last March. The aha! team continually seeks volunteers to help with revisions to the subject category tree and to assist with the placing of sites into that structure. The success of the aha! search engine depends on the completeness and accuracy of its entries. Browse through the aha! categories you know well and let the aha! team know what's missing. Web pages are added by suggestion—write aha@lanl.gov . The public version of aha! can be found at http://aha-public.lanl.gov .

Integrating Full-Text Search, the aha! Categorization Tool, News, and Tasks into the Administrative Resources (or Enterprise) Portal

Before the Administrative Resources Web page was created, office managers often had to go to three or four locations online to make purchases, submit travel requests, and the like, says Pat Hummer with Business Information Systems (CIC-13). "A lot of this information has been out there on the Web, but it's been a challenge for some to find," she said. Activated October 15, 1999, the Web site contains links to memoranda, job and hiring information, travel, purchasing, time and effort, computers, official documents, reports, and conference rooms in open and classified areas.

The design of the portal began with a needs assessment. To make the information useful to users, the development team collected data through an e-mail survey and telephone interviews. Nikki Goldman and Margaret Burgess of Communication Arts & Services (CIC-1) polled about 700 office managers, administrators, and secretaries asking what Lab Web sites they access regularly as part of their job; what information they are having trouble finding; what resources they have bookmarked to help them do their job; and what other types of administrative information they need, such as Department of Energy orders and administrative manuals. A common set of administrative "tasks" was identified and shortcuts to the pertinent information provided.

In addition, there is an administrative aha!. For this request, the aha! team expanded beyond categorizing science and technology information and worked with key people in support organizations to find information that Laboratory employees need about the environment, safety, health, organization, security, administrative, and personnel information. Usability tests with managers and administrative staff were conducted that resulted in modifications to the first design.

The page will be revised periodically to include new information useful to office administrators and managers. It can be accessed at http://int.lanl.gov/enterprise/ . Suggestions for improvements, resource categories, and other Web pages that can be linked to from this page can be sent to aha@lanl.gov by e-mail.

Next Steps

Meanwhile the IM Focus Team is now looking at the next steps for the administrative resources (enterprise) portal. The Gartner Group characterizes an enterprise portal as having index and search of multiple data repositories, structured and unstructured data, and internal and external data. It needs a taxonomy, application integration, security/permissioning, and personalization.

With aha!, we have begun to organize and structure "unstructured" content into an institutional taxonomy, adding the indexing and classification—the metadata that makes sense of it all. Librarians and aha! team members Mona Mosier and Kathryn Varjabedian would like to try semiautomatic slotting of content into categories in the future. Our ability to index, search, and categorize needs to evolve. We need tools to make it easier to administer, capture more useful content, and create a richer taxonomy.

And we also need to develop the tools that will facilitate interdisciplinary collaborations, allow us to identify access rights to information, and personalize our "desktop" to our needs and interests.

ESD Customer Survey Results

by Nikki Gaedecke, The Remote Electronic Desktop Support (REDI) Project, CIC-2 Desktop Group

Summary

From October 18 to November 1, 1999, the Electronic Software Distribution (ESD) Web site conducted its annual survey. The ESD team sent an e-mail to all ESD customers informing them of the online survey, and we also posted the survey under the Laboratory's "What's New" link on the internal Web site. Results were then tallied and published at the end of December 1999 on the ESD Web home page and the Lab's "What's New" link.

A summary of survey results include the following:

- received 776 completed surveys,
- received 397 open-ended comments,
- 80% of respondents are PC users,
- nearly 50% of respondents use ESD once a month or more,
- over 90% of respondents favorably rate ESD products and services, and
- over 80% of respondents favorably rate the getting and installing of software.

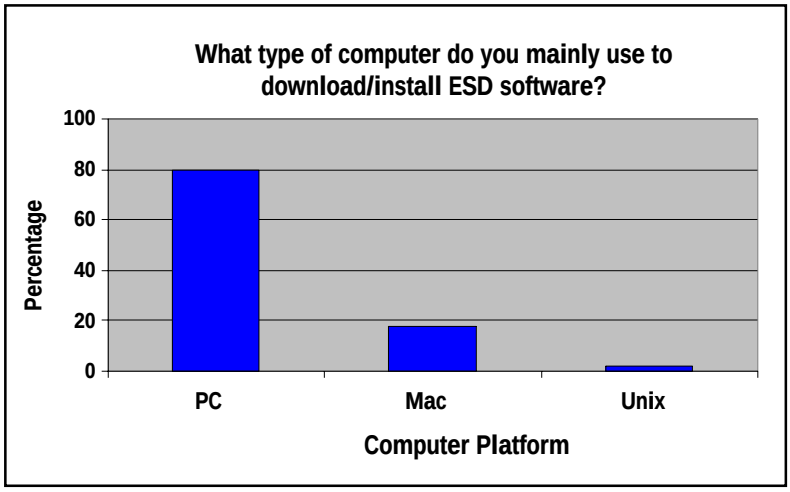
What is ESD?

ESD is the Electronic Software Distribution Web site that enables Laboratory personnel to purchase software licenses and to download and install the software applications. A convenient and cost-saving tool, ESD provides software at discounted prices through site-licensing and bulk acquisitions.

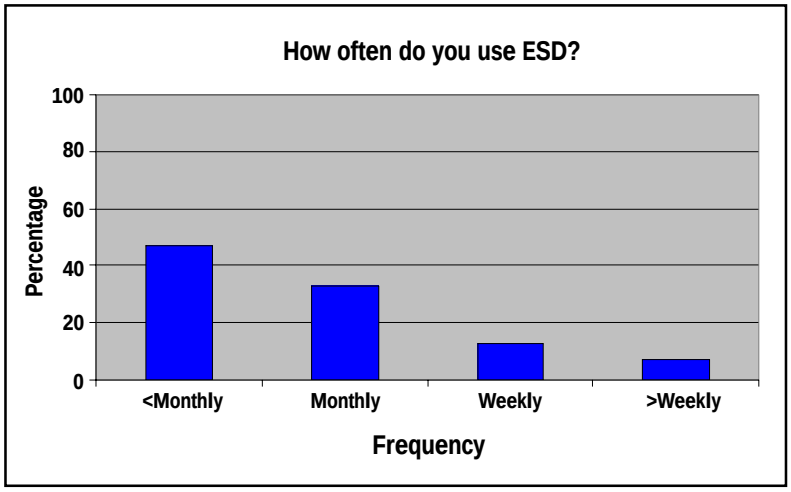
http://esd.lanl.gov

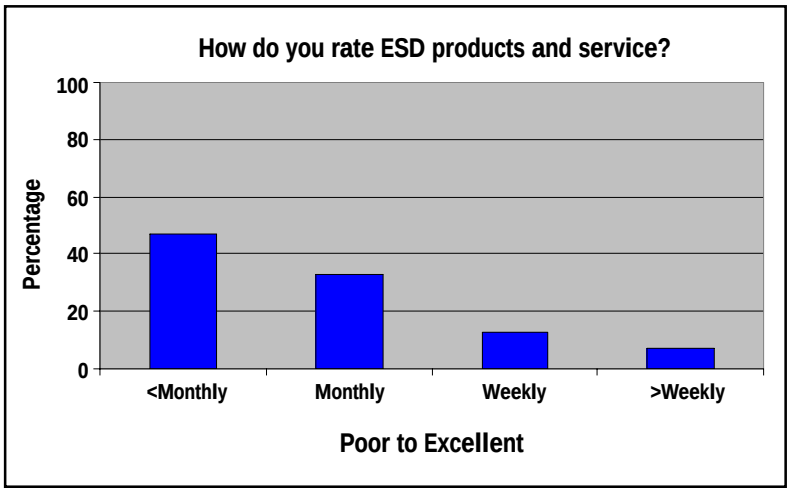
A Sampling of Summary Results

The following graphs show a sampling of survey results.



A Sampling of Summary Results—cont.





Open-Ended Comments

With the 397 received comments, we clustered them into various categories and have responses for the following headings. In special cases, we responded directly to individuals who provided Z numbers and had unique problems.

Net Installs

Because technology issues preclude having all products be made available as Web downloadable, the ESD team is unable to provide this option. In fact, software products that are "wrapped" for download and are 100MB are not reliable. Also, customers wanting to download a large software application would need more space than the installed application requires. For example, a customer wanting to download Microsoft Office 97 would need up to 1GB to accommodate the downloaded file, the install, and the actual application. The ESD team will look into other technologies as they develop.

Software Requests

The ESD team is always open to suggestions for software: simply e-mail esdtech@lanl.gov . For us to sell software, we must have enough interest to support a site license, and the software may not challenge an Information Architecture software standard. For example, ESD cannot offer WordPerfect because the current software standard is Microsoft Office 97/98.

Media and Documentation

As an electronic distributor of software, ESD does not sell media and documentation. As a courtesy, we link from the ESD Web site to multiple, online vendors who sell these products. Customers should direct any ordering questions to the appropriate vendor. See the following list for vendor contact information.

- Computer Corner: http://www.compcorner.com/lanlsoft/
- CJ Enterprises: http://206.124.177.43/LANL/LANLESD.asp
- Otowi Station: http://208.239.227.121/otowi/
- R Books: http://www.rbooks.com/compbooks.asp

Secure ESD

We are working on a secure ESD server system that we hope to offer in a couple of months.

Anti-Virus Assistance

After removing the "Guest Session" capability to enter ESD, many respondents complained because they had to log-in to ESD to access the Anti-Virus Assistance. In response, we placed an Anti-Virus link on the ESD entry page to circumvent log-in for individuals who download the monthly virus definition files.

Entry/Access

As a result of the limitation of HTML technology and our desire to meet the lowest common denominator of customers' Web browsers, we cannot program ESD entry with the Enter/Return key. Unfortunately, all users must click on the Enter button. However, we have set up "cookies" on the ESD Entry page that will remember your Z number.

Search (Find and Order)

Because of customer feedback, we have enabled a "search by category" option. In other words, when customers access the "find and order" option, they will be able to search by "databases" or "spreadsheets."

Complete Survey and Its Results

To view the complete survey results, which include the actual survey, responses presented in graphical format, and all received open-ended comments, go to http://esd.lanl.gov/esd/surveymain.html .

Here's a sampling of the respondents' comments.

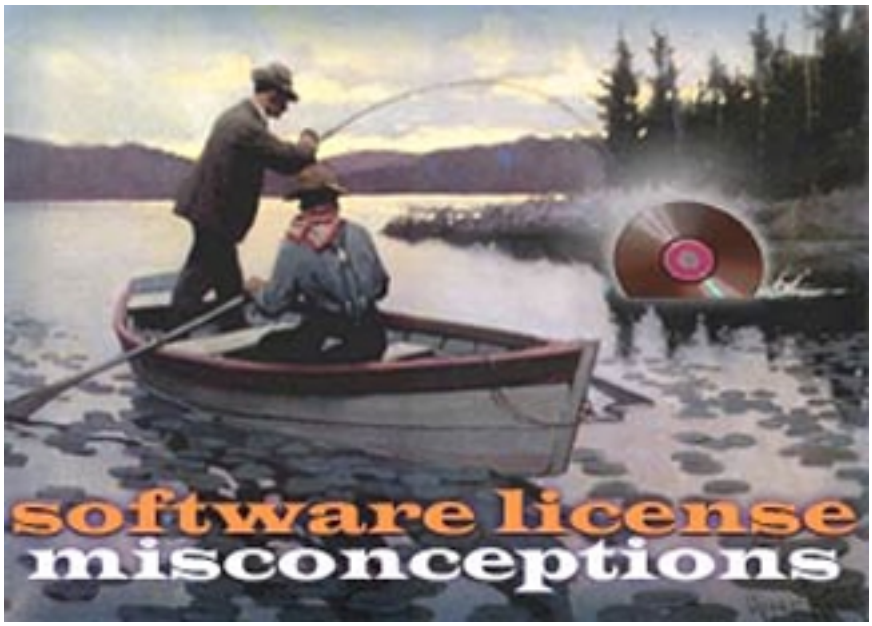
"I do like the site license and media-less approach better even with the greater complexity. Aside from being the fiscally responsible thing to do, it saves me from having to keep track of distribution media, people stealing my software application package (especially MS Office) from my office, and having to keep track of countless software licenses. All in all ESD is a very important service to the Laboratory."

"I've been with the Lab less than a year, and the ESD service is actually one of the things that is handled better than my previous employer, a high-tech computer manufacturer!"

"Can you add another screen that sorts products by their FUNCTION? When I needed a X-client I had to look at the description of EVERY product to find one. By function I mean categories like: compilers, office products, web browsers, x-emulators, anti-virus, web authoring, etc. Need to look for product both alphabetically and by function. Generally a very good service for the Lab."

Software Licensing Misconceptions

by Nikki Gaedecke, The Remote Electronic Desktop Support (REDI) Project, CIC-2 Desktop Group



What do fishing, driving, liquor, and software all have in common? What? Drunk men? Nope. Seriously, they require that you have a license.

Technically, you can go fishing without a license; drive without a license; sell liquor without a license; even use software without a license. BUT… they are all illegal actions resulting in penalties. Software piracy or software theft is the unauthorized duplication or use of computer software.

While ESD does not set or enforce software policy, we would like to remind people of their responsibilities. Many people have misconceptions about what a software license is and what the rules are. For example, a software license and a software application are not the same thing: a software application is a computer program; the software license gives the user authority to use the software application. Installing software with your coworker's CD-ROM is legitimate, but you would be running an illegal application unless you've purchased a license. On the other hand, you can purchase a single software license and reinstall the software multiple times on the same machine.

Laboratory Policy

In the Laboratory's *Administrative Manual*, it states that employees must follow commercial software restrictions. Failing to comply with software regulations, "may result in disciplinary action, up to and including termination" (AM 714.07—714.11: http://admin-manual.lanl.gov:1500/pdfs/adm/am714.pdf).

What Is a SoftwareLicense?

A license is a certificate that allows you to legally install and use software. Each software company has its own licensing policies, which can change from product to product. In general, to legally use software means that you:

- purchase an electronic license from ESD, or
- have the original media (floppy disks or CD) and documentation that comes with shrink-wrap software (boxed), or
- have a hard-copy "Certificate of Authentication" that typically comes with new computers/new operating systems.

While these guidelines are general, you should read the licensing agreements to definitively know the policy per license, per product. Most commercial software companies display their licensing agreements at the beginning of the software's installation; a screen appears asking users if they "accept" or "do not accept" the licensing agreement.

Using Software Illegally

- You "inherited" a coworker's computer with software already installed and you don't own the licenses.
- You borrowed your coworker's CD to install the software but haven't purchased a license.
- You bought one license for Adobe Photoshop but you have the application installed on two machines (e.g., home and office, or workstation and laptop).
- You change from one operating system to another (Macintosh to Windows), install Windows applications, and you do not purchase new software licenses.

- You purchase an upgrade Photoshop license on ESD but you don't own the full application Photoshop license.

Misconceptions about Software and Licensing: Do's and Don'ts

The Do's

- You must have a license for every software application you use.
- You must have a license for software on every machine (for example, licenses for home software)
- You must format your hard drive before giving a computer away as stated in the Computer Security handbook: http://int.lanl.gov/orgs/s/s5/handbook.shtml .
- You must own a previous license to purchase an upgrade license.

The Don'ts

- You cannot transfer Macintosh software licenses to Windows software licenses.
- You don't need to purchase another license if you need to reinstall the software.
- Don't be a software pirate; in other words, don't use software applications without purchasing licenses.

ESD Licensing

Regardless of how you install software on your computer, you must purchase a license to be legal. Unlike shrink-wrap software (boxed), software purchased electronically does not have a paper license. Instead, acquiring software electronically through ESD requires that you buy an electronic license through ESD.

One of many ESD benefits is the licensing management utility. By using either the "Your Licenses" option or the "License Utility" option, ESD users can generate a list of their ESD-purchased software licenses. ESD does not maintain records of software licenses purchased through other means such as vendors, other Web sites, preinstalled software with new computers, etc. While ESD sells software licenses, our customers have many choices to install the software: download from ESD, purchase CD from vendor, or borrow CD from co-worker (as long as you buy a license for the software).

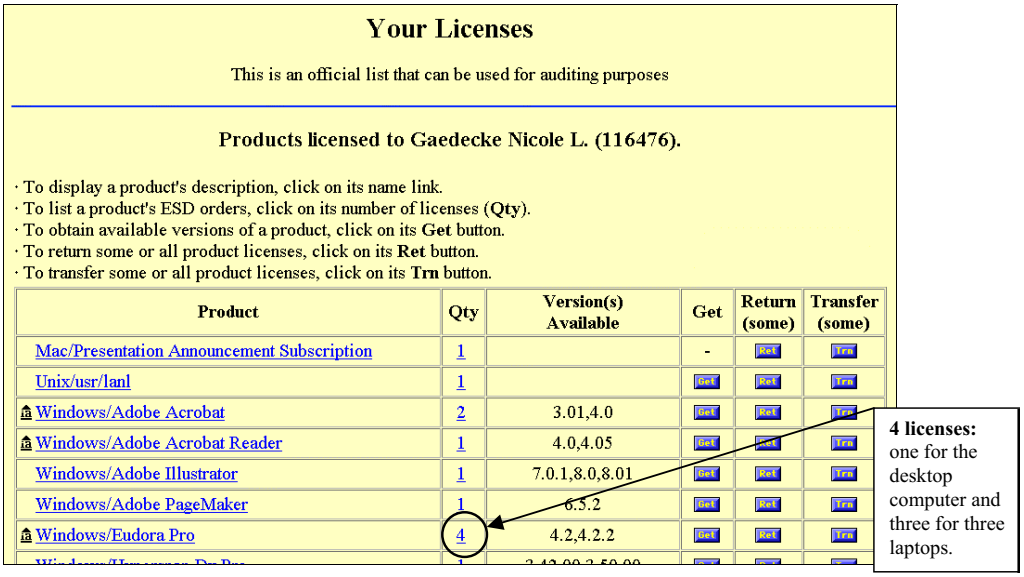
To View Your ESD Licenses:

1. Log on to ESD at http://esd.lanl.gov .
2. Click the **Your Licenses** button on the ESD main page. A list of licenses will appear.

Note: If you do not see a complete list of licenses, your organization may have "license custodians." These individuals purchase, for example, 100 copies of Adobe Photoshop licenses in their names on behalf of their 100 users. The license remains under the custodian's name, not the user's name. In addition, other reasons may be that you may have purchased the license elsewhere, or you simply failed to purchase a license. In the figure below, you can see a sample from the ESD license utility.

Questions

If you have questions about software licensing, contact your system administrator. For assistance with computer security, contact your organizational computer security representative (OCSR).



Your Licenses

This is an official list that can be used for auditing purposes

Products licensed to Gaedecke Nicole L. (116476).

· To display a product's description, click on its name link.
· To list a product's ESD orders, click on its number of licenses (**Qty**).
· To obtain available versions of a product, click on its **Get** button.
· To return some or all product licenses, click on its **Ret** button.
· To transfer some or all product licenses, click on its **Trn** button.

Product	Qty	Version(s) Available	Get	Return (some)	Transfer (some)
Mac/Presentation Announcement Subscription	1		-	Ret	Trn
Unix/usr/lanl	1		Get	Ret	Trn
Windows/Adobe Acrobat	2	3.01,4.0	Get	Ret	Trn
Windows/Adobe Acrobat Reader	1	4.0,4.05	Get	Ret	Trn
Windows/Adobe Illustrator	1	7.0.1,8.0,8.01	Get	Ret	Trn
Windows/Adobe PageMaker	1	6.5.2	Get	Ret	Trn
Windows/Eudora Pro	4	4.2,4.2.2	Get	Ret	Trn

4 licenses: one for the desktop computer and three for three laptops.

Moving an NT Workstation from a Workgroup to a Domain

by Andy Ryan, Microsoft Certified Systems Engineer, EES-13, Integrated Geosciences Group

Microsoft Workgroups and Domains

A networked NT workstation must be a member of a workgroup or a domain. In a workgroup environment, all user accounts reside on the workstations and resources are shared using peer-to-peer networking. Since workgroup security or authentication does not exist, all that is needed to join a workgroup is the workgroup name. Hence, joining a workgroup is technically meaningless except that the workstation NetBIOS name shows up in the browse list under the workgroup name. A workgroup environment is acceptable when the number of workstations to be managed is small. Microsoft recommends a workgroup with 20 workstations or less.

Conversely, in a domain environment, accounts are centralized on domain controllers. This allows administration of accounts from a single location. Domain security is robust and handles both user and machine accounts. By default, only domain administrators and account operators can add workstations to a domain.

Moving from a workgroup to a domain is a three-part process. First, NetBIOS connectivity must be established from the workstation to the primary domain controller (PDC). Second, the workstation must join the domain. Third, and optionally, the user's profile should be migrated. Microsoft might have provided a bundled migration utility to handle these tasks, but has not. Thus the necessity of this article.

For the remainder of this discussion, a TCP/IP-only environment is assumed and the workstation and PDC exist on different subnets. For illustration, the domain name is MYDOM, the domain user account is NEWACCT, the workstation user account is OLDACCT, and the workstation NetBIOS name is WS1.

Establishing NetBIOS Connectivity

Before attempting to add a workstation to a domain, make sure that the workstation can see the PDC via NetBIOS. At this point, as a prudent NT administrator, you should have WINS, DNS, or LMHOSTS implemented on your network. If you are relying on broadcast for NetBIOS name resolution, you really should consider WINS.

The following steps are performed from the workstation. Please refer to Fig. 1 for the following discussion. Initially, to verify connectivity to the PDC from the workstation, examine the machine names of the PDC. If the NetBIOS name differs from the TCP/IP name, verification is simple. This is the only situation I can think of where having dissimilar machine names is advantageous. If the PDC machine names are different, ping the NetBIOS name of the PDC. If the ping is successful, connectivity is verified and this part is complete.

Otherwise, if the NetBIOS and TCP/IP names of the PDC match, first ping the IP address of the PDC. If the ping is unsuccessful, the PDC is down or unreachable. If the ping is successful, type "NET VIEW \\PDCNAME" at a command prompt, using the PDC's NetBIOS name. If you get "Access Denied," connectivity is verified. When "Access Denied" is returned from the NET VIEW command, you have reached the PDC with your request and the PDC has challenged you for credentials. If NET VIEW returns "The network path was not found," you are not getting NetBIOS name resolution.

To troubleshoot a NetBIOS name resolution problem, first enable LMHOSTS lookup on the workstation and create an LMHOSTS file that has an entry for the PDC using the #PRE and #DOM switches. The LMHOSTS file could simply consist of the following line:

128.165.xxx.xxx PDCNAME #PRE #DOM:MYDOM

The LMHOSTS file must reside on the workstation in %SYSTEMROOT%\SYSTEM32\DRIVERS\ETC and be named LMHOSTS. Now type NBTSTAT –R at a command prompt to purge and reload the name cache. Then type NBTSTAT–c to verify that the PDC is present in the name cache. Note that

the switches to the NBTSTAT command are case sensitive. If the cache is empty, LMHOSTS lookup is not enabled or the #PRE switch is not present. Now retry the "NET VIEW \\PDCNAME" command. You should get "Access Denied." If you do, your WINS or DNS isn't working. Alternatively, if you still get "The network path was not found," you have a genuine anomaly on your hands. Things like subtle typos in the LMHOSTS file or a faulty or misconfigured network could be the cause.

Joining the Domain

With NetBIOS connectivity established between the workstation and PDC, joining the domain is easy. First, login to the workstation as the local administrator, open the Network applet, and click "Change." Then click "Domain" and enter the domain name. Next, check "Create a Computer Account in the Domain" and enter the credentials of a domain administrator or account operator. Click "OK" and you should see the "Welcome to the MYDOM Domain" window. By checking "Create a Computer Account in the Domain," a machine account is automatically created on the domain and the Domain Admins global group is added to the Administrators local group on the workstation.

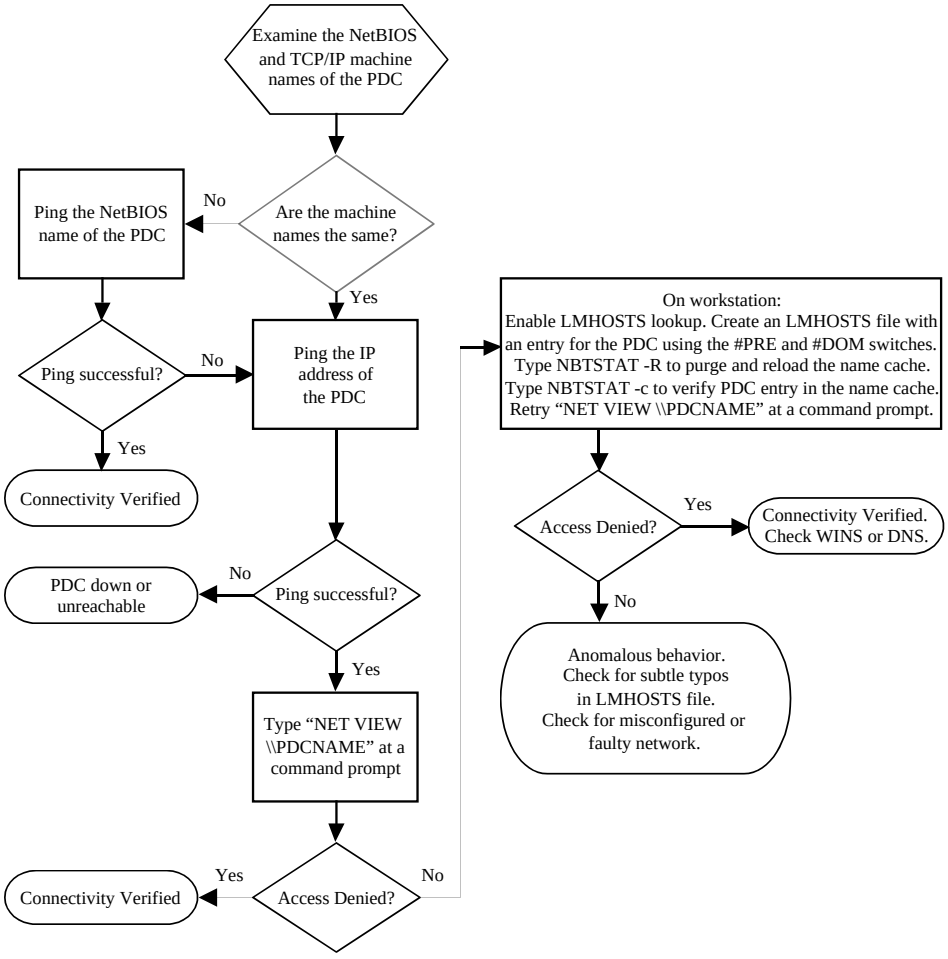


Fig. 1. Verfying NetBIOS Connectivity from a Nonmember Workstation to the PDC.

To finish this part, login to the domain from the workstation using a domain admin or account operator account to verify the successful creation of the domain machine account. You may need to synchronize the domain if you receive "The trust relationship between this workstation and the primary domain failed" when trying to login. Note that you can bypass the domain login and still do a workstation login by choosing the machine name, WS1, instead of the domain name, MYDOM, in the "From" drop down box.

Migrating the Profile

The reason the desktop profile should be migrated is that the user will be assigned a different account that resides on a domain. This new account starts with the default profile for that workstation, so the user's desktop will probably look nothing like it did before. At this point, you could simply create a domain account for the user and rebuild the desktop by manually creating shortcuts and such, but migration of the old profile will provide a quick and transparent move to the domain environment. The following steps are derived from trial and error and may well be improved upon, but this method works.

1. Create the domain account, NEWACCT.
2. Make sure NT SP4 or later is installed on the workstation.
3. Login to the workstation as MYDOM\NEWACCT. This step creates the new profile directory containing the default profile. Restart Windows.
4. Note: After the first login, the profile will be located in %USERPROFILE%, which will probably be C:\WINNT\PROFILES\NEWACCT. Importantly, if you name the domain account the same as the workstation account, which is a very reasonable thing to do, Windows appends a ".000" to the profile directory. In that case, the new profile would be stored in C:\WINNT\PROFILES\ NEWACCT.000. Thus, the workstation account profile would reside in C:\WINNT\PROFILES\NEWACCT and the domain account profile would be located in C:\WINNT\PROFILES\NEWACCT.000. From hereon, we will assume the account names are different and the ".000" was not used.
5. Login to the workstation as the local administrator and delete everything in C:\WINNT\PROFILES\NEWACCT, leaving only the directory name intact.
6. Run the System applet and click the "User Profiles" tab. Highlight WS1\OLDACCT and click "Copy To." Next, under "Copy profile to," enter C:\WINNT\PROFILES\NEWACCT. Now click "Change," "Show Users," and choose MYDOM\NEWACCT. Click "Add," then "OK," then "OK" again in the "Copy To" window, and "OK" again in the "System Properties" window.
7. You may now opt to run User Manager and add MYDOM\NEWACCT to the local administrators group. Also, the local accounts and profiles may be deleted for security purposes. You cannot delete Guest or Administrator. It is recommended to disable the Guest account and set the password of the Administrator account to a tough password. You should keep track of these tough password(s) because the local Administrator account provides a way to administer the computer should the domain fail.
8. Restart Windows and login as MYDOM\NEWACCT. You should see the desktop of WS1\OLDACCT.

Summary

Migrating NT Workstations from workgroups to domains is fairly complex and labor intensive because of the lack of a migration utility. There are three major parts involved in the migration. These are verifying NetBIOS connectivity from the workstation to the PDC, adding the workstation to the domain, and migrating the user profile.

To accomplish the first part, use the methods described in Fig. 1 to simplify the process of verifying NetBIOS connectivity from the workstation to the PDC. The second and easiest part is adding the workstation to the domain. The reason the second part is easy is because the odds of success are high as a result of the completion of part one. The third and last part, migrating the user profile, can be tricky, but the procedure makes sense after a couple of migrations.

An Update on Computer Security Initiatives

by Don Willerton, Technical Staff Member, Information Security Management (ISecM) Project Office, CIC Division

It was about a year ago that computer security at LANL suddenly took on national interest. You might remember the news of a Chinese nuclear weapon design that looked similar to ours, the revelation of the FBI investigation of Wen Ho Lee, and the sudden, two-day stand-down for computer security and counterintelligence training. What ever happened to the security initiatives that were started? Does the Nine-Point Plan still exist? What came of all those investigations and audits?

Who Said We Have to Do What?

Rest assured that the Nine-Point Plan is still alive and kicking, with the addition of several other security guidance documents, mandates, and corrective actions. And we've had a few more DOE orders added to help us with our cyber security. Let's look at an overview of what's happened, the status of the various initiatives, and where we still need to go.

The Cyber Security Component of the "Go Green" Legacy—Mandated before April 1999

As a yearly evaluation, the Laboratory has had for several years a "report card" that gave a red (bad), yellow (not-so-bad), and green (okay) designation to a number of security areas and issues, including guards, guns, fences, information, and documents, in addition to computer security. A major initiative was undertaken within the last two years to "Go Green" across all of the boxes of the report card. Last August the DOE Office of Oversight and Assessments (OA) performed the latest full audit for the report-card evaluation. Computer security, judged yellow, was the one item that lagged the others in being judged green, and the reasons are cited below.

The Nine-Point Plan—Mandated April 14, 1999

In mid-April, a major computer security initiative was agreed upon between the Secretary of Energy and the three weapons Laboratories. There were nine major areas of work, each being further defined by milestones and performance measures, including some immediate, medium-term, and long-term actions. The Laboratory's stand-down for security-immersion training and the creation of a cyber security policy board are examples of immediate actions. Some medium-term actions included prevention of accidental transfer of classified data and the audit of unclassified computers and e-mail. Some long-term actions were creating a three-level security network model, eliminating nonvolatile memory from desktops, and making it physically impossible to connect to the wrong network.

The Six Enhancements—Mandated May 1999

In mid-May, six additional cyber security "enhancements" to the Nine-Point Plan appeared on the DOE Web site. They addressed such things as continuous monitoring of cyber system security, DOE-provided training for system administrators, better use of technology in secure environments, and requirements for stringent control over classified ports, cabling, memory, and media. Some DOE officials interpreted the six enhancements to apply only to DOE and not to its contractors. Other DOE officials interpreted otherwise, which seems to be the interpretation that we get to live with.

The Information Security Management (ISecM) Project—Mandated November 1999

In May, the DOE Defense Programs manager chartered a special team of people from the three Laboratories and some of the plants, to develop a long-term security strategy. This strategy was intended to bring the entire DOE Complex (labs, plants, and headquarters) to a state of cyber security "preeminence" against the insider threat and "world class" against all other cyber threats. This became known as the Summer Study, and produced a report the last of August 1999. It was followed by an implementation plan in October, with the intention that it would be funded and implemented over a three-year period.

The OA Findings—Mandated August 1999

In August, an audit by OA identified four cyber security-related issues addressing the approval process for Foreign National access to cyber-systems, the process for identifying and correcting system vulnerabilities on the unclassified/open (informally, "yellow") network, and general processes for control of effective computer security.

The DOE Office of Counterintelligence Audit Findings—Mandated September 1999

In August, a team of auditors from the DOE Office of Counterintelligence came to the Laboratory, resulting in three cyber security findings. These addressed the inclusion of the Information Security Office in the security incidents and the process for granting sensitive-country Foreign Nationals access to computer systems, and issues relating to system administrators.

DOE Order 205.2—Mandated November 1999

In November, General Habiger, DOE Director of the Office of Security and Emergency Operations, issued DOE Order 205.2, giving official guidelines for the approval of access to all LANL cyber systems for sensitive-country Foreign Nationals. The guidelines are posted on the Web at http://s6server.lanl.gov/isec/fva.html or see the article "Policy Changes for Foreign Visits and Assignments" in BITS January 2000 issue.

Draft DOE Order 205.3—Mandated November 29, 1999

Also in November, the DOE Chief Information Officer, John Gilligan, issued a new password policy.

The Status of the Mandates

It's a little difficult to cover the progress made in responding to each point of each of the mandates. Let me summarize the major differences between last year and the present.

- In December 1999, another visit by OA gave the entire report card a green designation. This was a significant achievement for the Laboratory and all of the concerned parties should be recognized for working long and hard to make this happen. For the cyber security component, the green designation was mostly due to aggressive and persistent scanning of the yellow network for system vulnerabilities done between September and December.
- The Nine Point Plan is the continuing, dominating mandate, having the most milestones and measures. Good progress has been made. Currently, forty-eight milestones have been met, out of a total of fifty-five; twenty-four performance measures have been met out of a total of thirty. Milestones and performance measures have been completely satisfied in three of the nine points (except where they are ongoing). Several of the concerns of the Nine-Point Plan also address the Six Enhancements.
- In particular, the secure (informally, "red") network is being changed from a "shared" network topology to a "switched" network topology. This is a fundamental, and expensive, change that prevents a host of network intrusion vulnerabilities. Completion is targeted for September 2000. Along with this, we are developing the capability for identifying ports with the particular machines plugged into them, taking away the threat of subversive use of network connections.
- Several options for answering the problems of removable media, local storage, the two-person rule, and other security considerations have been identified, explored, and demonstrated. Various options include the networking of stand-alone workstations, use of the keyboard-video-mouse (KVM) technology, and the use of a secure version of the LANL Desktop-On-Demand (thin client) software.
- The ISecM recommendations described a significant change in security, network, and information architecture across the DOE Complex. The recommendations, as stated, cost several hundreds of millions of dollars, which no one had. There were a few iterations that proposed lesser changes, with subsequently lower price tags, but with no acceptance. Ultimately, the ISecM project and team were placed into "stand-down" mode when zero dollars were found to fund the recommendations.
- The Foreign National Requirements Working Group, a committee of the LANL Information Policy Board (IPB), has developed a LANL draft policy that addresses the concerns of permitting Foreign National access to LANL cyber systems. Also, an automated version of the process for Foreign National Visits and Assignees (Form 982) has been developed and is close to production. This addresses both the OA audit finding and DOE Order 205.2.
- The Laboratory took on an institutional overhaul of passwords, developing and instituting a new policy with a short turnaround time for Lab-wide compliance. Our password policy is very close to the guidelines put forth by DOE Order 205.3.

What Now?

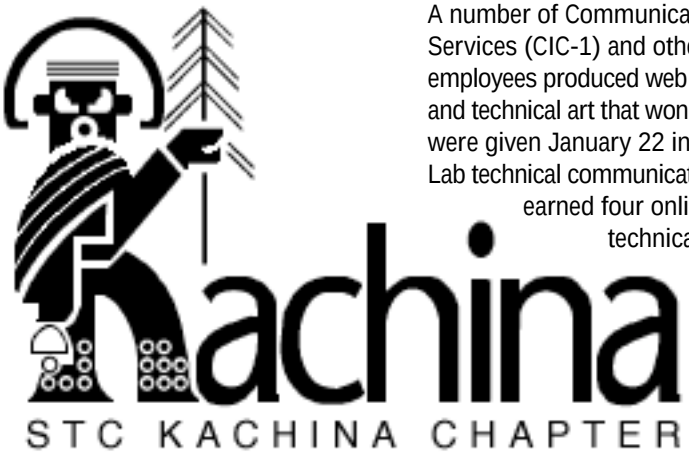
In the last year, significant strides have been taken to meet all of the above mandates and to move the Laboratory, as an institution, into a greater, more measurable and more accountable security environment. Unfortunately, in the eyes of an auditor, probably not one of the mandates has been completely satisfied.

- Even "Going Green" is a point-in-time measure. The scanning performed was at a rudimentary level of automated system investigation. Over the next few months, the scanning parameters will be adjusted to seek out and identify more varieties of system vulnerabilities, with the intent of being able to periodically scan at the most comprehensive level.
- The Nine-Point Plan activities have addressed and satisfied most of the original requirements and intents. The ones left are the very hard ones. Not only hard, but also costly (several tens of millions of dollars). In fact, some of the original Nine Point Plan objectives may not be possible for the LANL environment without a major program that makes substantial changes to LANL and the DOE Complex, such as what ISecM proposed. One response to this may be to separate some of the incomplete items into a new initiative and close out the Nine-Point Plan as best as possible without them.
- With the lack of funding for widespread Laboratory initiatives, much of the responsibility and cost for remediating secure workstations and networks to the newer security standards falls directly on the divisions and groups. Given the cost for implementing solutions, some of the Nine-Point Plan initiative fulfillment have developed considerably longer timeframes than the original October 2000 finish date.
- The issue of Foreign National access is complex. The policy expresses a solution for a particular aspect of the issue, but it has not withstood the inspection and review of outside audits or other parties with vested interest. There may be changes yet to come.

Summary

We are in better shape than we've ever been. Cyber security is taken more seriously at LANL than it has ever been. Along with the heightened awareness and accountability, though, has come a deeper appreciation for answering cyber threats with real, substantial, and complete answers. Consequently, the price continues to go up and any additional funding seems but a vapor in the wind. It may seem an intractable problem, but it may be that this reality will force different, but secure enough, solutions.

BITS Publications Earn Society for Technical Communication Awards



by Denise Sessions, BITS Managing Editor, CIC-1 Communication Arts & Services

Los Alamos National Laboratory publications received five Distinguished Technical Communication awards, while more than two dozen other communication products also earned awards from the Society for Technical Communication's (STC's) Kachina Chapter.

Each year STC sponsors competitions for technical publications, technical art, and online communication products, such as World Wide Web pages. This year, the Kachina Chapter of New Mexico received more than 130 entries from six southwestern and western states. In judging them, the chapter exchanged entries with those received by the East Tennessee's STC chapter. Such a judging exchange eliminates potential conflicts of interest when judges submit their own work in the competitions.

A number of Communication Arts and Services (CIC-1) and other Laboratory employees produced web sites, reports, and technical art that won awards, which were given January 22 in Santa Fe. Lab technical communications products earned four online awards, 17 technical art awards, and 14 technical publication awards. Three levels of awards are given: Distinguished Technical Communication, Excellence, and Merit.

BITS received three Awards of Merit:

- Technical Art category for the cover of *BITS Introduction to Computing at Los Alamos, 4th ed.* Chris Lindberg of CIC-1 designed the cover.
- Online Communication category for the online version of the bimonthly *BITS Computing and Communication News*. David Van Etten of CIC-1 programmed the HTML for the September 1999 issue submitted for competition.
- Technical Publications category for the hard-copy version of the bimonthly *BITS Computing and Communication News*. Julie Medina, Donald Montoya, and Denise Sessions, all of CIC-1, produced the three consecutive issues (June/July '99, August '99, and September '99) submitted for competition.

Entries that won Distinguished Technical Communication awards are

"For the Seventh Generation," technical publication

"Physics Division Progress Report, Jan. 1, 1997–Dec. 31, 1998," technical publication

"Breast Cancer Awareness Poster," technical art

"Women: Putting Our Stamp on America," technical art

"DARHT: Dual Axis Radiographic Hydrodynamic Test Facility," technical art

A complete list of the award winners can be found online at http://www.lanl.gov/cic1/99-00_stc.html .

STC's International Competitions

The five entries that won Distinguished Technical Communications awards have been submitted to the STC's international competitions; judging for the internal competitions will take place during February. The winners from competition will be displayed at the International STC Conference in May in Orlando, Florida.

About STC

The Society's members include writers, editors, illustrators, printers, publishers, educators, students, engineers, and scientists employed in a variety of technological fields. With more than 20,000 members worldwide, STC is the largest professional organization serving the technical communication profession. The central/northern New Mexico chapter has over 100 members. For more information about the Kachina Chapter, see http://www.stc.org/region5/nkc/ .

CIC Division Undergoes Review in March

by Denise Sessions, BITS Managing Editor, CIC-1 Communication Arts & Services

The CIC Division Review Committee (DRC) will meet March 27–29, 2000, in Los Alamos.

The DRC committee members are appointed by the Laboratory Director to review Division activities and to advise Division and Laboratory management regarding CIC technical and operational activities. The committee is responsible for assisting the Division in maintaining excellence and leadership in computational science, information management, and communications. They are charged with identifying new directions and opportunities appropriate for the Division, advising the Division on its research and development activities, forging collaborative ties with universities and other laboratories, and helping the Division sustain a high-quality, stimulating working environment. The DRC reviews the Division's organizational design and advises on its effectiveness for its customers.

The following biographical summaries of the DRC members will give you some familiarity of their current positions, education and expertise, and background.

David M. Cooper, *Associate Director, Computation Directorate, Lawrence Livermore National Laboratory (LLNL)*

Education and Expertise: Dr. Cooper received his B.S. degree in physics from Texas Technological University in 1962, his M.S. degree in aeronautics and astronautics from Stanford University in 1965, and his Ph.D. in physics from York University in 1974. His expertise is in system software and advanced algorithm development.

Background: In his present position, Dr. Cooper provides technical leadership for the direction, execution, and development of a broad research and support program in computational science and advanced information technologies. He is responsible for the management, maintenance, and development of the Laboratory's scientific computational centers. He supervises the operation of the LLNL Accelerated Strategic Computing Initiative computers and problem-solving environment, which will provide the computational underpinnings for the Science-Based Stockpile Stewardship program. He is responsible for the coordination and management of the government-industry collaboration on high-speed storage systems. He is responsible for the operation, maintenance and development of classified and unclassified networks. He serves as the CIO for LLNL.

Dr. Cooper has served on these advisory committees: External Advisory Committee for Center for Research on Parallel Computation; External Advisory Committee for Florida A&M Department of Science and Engineering; Advisory Board for Northern Arizona University Department of Engineering; Advisory Board for San Jose State University Department of Engineering; and Presidential Advisory Committee on High-Performance Computing and Communications, Information Technology and the Next Generation Internet.

Stuart Feldman*, Director, IBM Center for Advanced Commerce, Department Group Manager, Networked Computing Software Research*

Education and Expertise: Dr. Feldman received his A.B. in astrophysical sciences from Princeton University (magna cum laude) in 1968 and his Ph.D. in applied mathematics from Massachusetts Institute of Technology in 1973. His expertise includes Internet services and applications management, and software support of advanced networks, software architectures/ configuration management/development environments.

Background: Dr. Feldman has taught courses in advanced topics in software systems, development environments for large scientific systems, and introduction to applied mathematics. He has authored numerous publications. Some of the boards and committees that Dr. Feldman serves on are Steering Board, Joint ACM/IEEE-CS Task Force on Software Engineering as a Profession; National Council on Science and Technology Education; Numerical Algorithms Group (Oxford, UK) Technical Policy Committee; screening panel for NSF Infrastructure Grants; and final advisory panel for NSF CER grants.

***Michael A. Harrison**, Professor Emeritus, Computer Science Department, University of California, Berkeley*

Education and Expertise: Dr. Harrison received his B.S. degree in electrical engineering from Case Institute of Technology in 1958, his M.S. degree in electrical engineering from Case Institute of Technology in 1959, and his Ph.D. in communication sciences from University of Michigan in 1963. His expertise includes multimedia systems, software environments, computer security, and theoretical computer science.

Background: In a previous career, Dr. Harrison started a software company, Gain Technology, to build a comprehensive multimedia authoring system for UNIX workstations. This product, called GainMomemtum, won the SunWorld Prize for the Best Product of 1992. He has authored seven books and over 200 technical publications in a variety of areas. His research interests began with switching and automata theory and shifted, at Berkeley, towards automata theory and discrete systems. This led to work on formal language theory and ultimately its application to fast parsing methods for compilers as well as getting the fastest practical algorithm for general context-free recognition. Switching to software research, a simple abstract model of protection in operating systems was invented. While on sabbatical at Stanford University, he helped build an advanced environment for electronic document production. Harrison is currently working with a colleague to develop the Ensemble System—a framework for the integrated support of interactive development of complex natural language and formal language documents.

***Kim Molvig**, Associate Professor of Nuclear Engineering, Massachusetts Institute of Technology*

Education and Expertise: Dr. Molvig received his B.S. degree in engineering physics from Cornel University in 1970 and his Ph.D. in physics from the University of California in 1975. His expertise is in digital physics—a new computational method for solving fluid dynamics.

Background: Dr. Molvig teaches courses in introductory plasma physics, plasma kinetic theory, plasma transport phenomena, nuclear physics, and seminars on fusion to undergraduate and doctoral students. His research area includes theoretical research and publication on controlled thermonuclear fusion. In 1998 he discovered digital physics which is a new form of scientific computing that, at least for fluid dynamics, solves the fundamental incompatibility of the laws of physics with the operations of computers. He founded and served as chief technical officer and director for Exa Corporation to commercialize digital physics. The company now has a breakthrough technology imbedded in a pure software product that is being adopted by major corporations around the world. He serves as the chairperson for the External Review Committee for the Applied Physics Division at Los Alamos National Laboratory. He has authored numerous publications and holds five U.S. patents.

***Cherri M. Pancake,** Associate Professor and Intel Faculty Fellow, Department of Computer Science, Oregon State University*

Education and Expertise: Dr. Pancake received her B.S. degree with distinction in design and environmental analysis from Cornell University in 1971 and her Ph.D. in computer engineering from Auburn University in 1986. Her expertise includes computer engineering and software support for high-performance computing.

Background: In her current career, Dr. Pancake has applied her ethnographic training to the problem of how tools interfaces might better support users' conceptual models, particularly in the application of visual techniques to parallel debuggers and performance analysis tools. Focusing on the needs of the scientific and technical programming communities, she has developed new visual idioms for a range of trace-based tools and breakpoint-style debuggers. With the advent of Web technology, she has extended that work into the area of network-based training and support for high-performance computing. Over a period of nine years, she served as principal or coprincipal investigator on contracts and grants to develop software tools on behalf of the U.S. Army Missile Command, NASA, Rome Air Development Center, IBM, and Intel Supercomputer Systems. At Auburn University and as a Visiting Scientist at the Cornell Theory Center, she was also active in a number of pilot studies conducted for industrial sponsors (including IBM Enterprise Systems Division, Supercomputer Systems Incorporated, Kendall Square Research, Intel Supercomputer Systems Division). Her previous career involved extensive ethnographic fieldwork, where she applied cross-cultural survey and interviewing techniques to study social change in Guatemalan Indian communities. She also serves as an advisor to industry on human factors engineering and user-oriented design.

Paul R. Woodward, *Professor of Astronomy and Fellow of the Supercomputer Institute, University of Minnesota; Director, Laboratory for Computational Science and Engineering; Director of Graphics and Visualization for the Army High-Performance Computing Research Center (AHPCRC); and Coordinator, Computational Continuum Mechanics Program for the AHPCRC.*

Education and Expertise: Dr. Woodward received his B.A. degree in math and physics from Cornell University in 1967 and a Ph.D. in physics from the University of California at Berkeley in 1973. His expertise includes difference methods, scientific visualization, grand challenge computation, theoretical astrophysics, and illustrative Web documents and images.

Background: Dr. Woodward was awarded a National Merit Scholar, a Woodrow Wilson Fellowship, and a Whiting Fellowship. He is a Fellow of Minnesota Supercomputer Institute. He was elected to the New York Academy of Science and won the Sidney Fernbach Award in High-Performance Computing from the IEEE Computer Society. He is an associate editor for Scientific Visualization, IEEE Computational Science and Engineering. Some of the committees and boards that Dr. Woodward serves on are: Advisory Board, National Center for Supercomputing Applications; Peer Review Board, National Center for Supercomputing Applications and Pittsburgh Supercomputing Center; External Advisory Committee, Center for Research on Parallel Computation; Advisory Committee, Minnesota Supercomputer Institute; Physics and Space Technology Advisory Committee, LLNL; External Advisory Committee, Physics Department, LLNL; CISE Advisory Committee, NSF Directorate for Computer and Information Sciences and Engineering; and many more.

Katherine Yelick*, Associate Professor, Computer Science Division, University of California at Berkeley.*

Education and Expertise: Dr. Yelick received her B.S. degree in electrical engineering/computer science from the Massachusetts Institute of Technology (MIT) in 1985, her M.S. degree in electrical engineering/computer science from MIT in 1985, and her Ph.D. in electrical engineering/computer science from MIT in 1990. Her expertise includes parallel computing, programming languages and compilers, and symbolic computation.

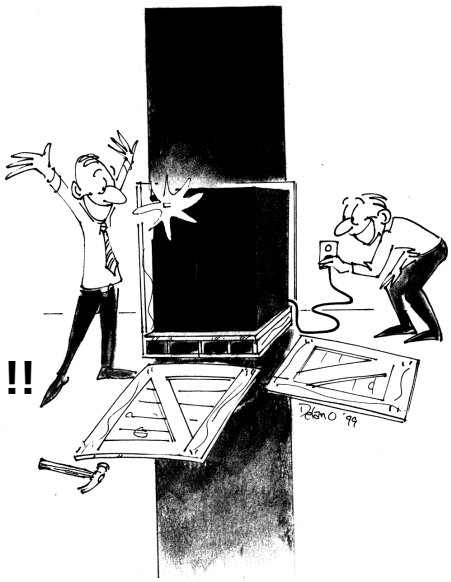
Background: Dr. Yelick currently teaches data structures and advanced programming and design of programming languages. Current research projects are: Titanium—compiler optimizations for explicitly parallel programs; Intelligent RAM (IRAM)—software support for IRAM machine; Split-C—a performance programming language for parallel machines; Multipol—a data structure library for distributed memory multiprocessors; Coordination in Parallel Programs—software support for coordinating parallel computation; and The Castle Project—integrated software support for parallel computing.

LOS ALAMOS NATIONAL LABORATORY

Research Library

http://lib-www.lanl.gov

The LANL Research Library offers a variety of training opportunities for the Laboratory community. Available sessions focus on specialized library databases and other electronic resources. A complete list of course offerings can be found at http://lib-www.lanl.gov/libinfo/training.htm. All sessions are available to individuals or groups at the library or your site. Arrange for a session by contacting the Library, phone 7-4175 or e-mail library@lanl.gov. Library tours are available on a drop-in basis every Wednesday at 1:00 p.m.



Attention Application Developers!!!

Need User Guides?

Need Online Help?

CIC-1's Software Documentation Team can help you improve your product by developing user manuals, online help, quick reference cards, or other documentation to help your users learn the application. Contact Sheila Molony, 5-1585, or sheila.molony@lanl.gov, for more information.

Computer Training

The Customer Service Group (CIC-6) offers technical computer training (enterprise information applications, communications, office administration, and Web authoring) and advanced computer training (programming languages, system administration, and advanced applications). To register for a course access our Web page at http://www.lanl.gov/internal/training/training.html . Or from the LANL home page select the links: Training, Computer. For further information about technical computer training call (505) 667-9559, and for advanced technical computer training call (505) 667-9399.

Technical and Advanced Technical Computer Training Courses

Communications	Office Skills 2000	Web Authoring and Browsing
• Eudora 4.2 • Meeting Maker 5.5.3	• Office Skills 2000–LANL Computing • Office Skills 2000–Professional Development	• Dreamweaver 3.0 • FrontPage 2000 • HTML
Enterprise Information Applications (EIA)	**Other EIA Courses**	**System Administration Training**
• Date Warehouse–Basics • Date Warehouse–EDS Reports • Data Warehouse–Passport • EDS–Basics • EDS–GUI • EDS–Training Plans • Infomaker • Invoice Approval System • Procurement Desktop • Purchase Card System • Recharge • Time & Effort GUI • Travel Foreign • Travel Domestic GUI • Web JIT	• Automated Chemical Information System (ACIS) • Directory Information System (DIS) • Financial Management Information System (FMIS) • Key/Core • Property Accounting, Inventory, and Reporting System (PAIRS) • Resource Planning Module (RPM) • Salary Management Tool (SMT) • Secretarial/Contract Service (SE) • Signature Authority System (SAS)	• IRIX (SGI) System Administration (Beginning) • IRIX (SGI) System Administration (Advanced) • IRIX (SGI) Network Administration • Linux System and Network Administration • Solaris 7 System Administration I • Solaris 7 Network Administration II • Solaris TCP/IP Network Administration • Unix and Widows NT Integration • Windows NT/2000 Optimization and Troubleshooting • Windows NT/2000 Workstation and Server • Windows 2000 Update for IT Professionals

Technical and Advanced Technical Computer Training Courses—cont.

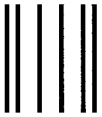
Programming Training	Application Training
• Advanced Perl Programming with CGI • ANSI/ISO C++ Programming Clinic (Advanced C++) • C Programming (Beginning) • C++ for Experienced C Programmers • Distributed Objects Using Corba • Distributed Programming with Java • Fortran 90/95 • Java Programming • Java Program Workshop • Javascript: Building Interactive Web Sites • Object-Oriented Analysis and Design • Perl Programming • Programming for Beginners Using Java • Shell Programming	• Advanced WWW Development • Apache Web Server • C++ and the Unified Modeling Language • Digital Visual Fortran • Foundations of IDL Programming • FrameMaker Basic and Advanced • IDL 5.0 Graphic Object Workshop • Oracle Classes — Advanced PL/SQL — Develop PL/SQL Program Units — Enterprise DBA Part 2: Performance Tuning — Introduction to Oracle for Experienced SQL Users — Oracle Developer Release 6: Build Forms I — Oracle Developer Release 6: Build Forms II — Oracle Developer Release 6: Build Reports — Oracle Developer Server: Deploy Web-Based Applications • Sendmail–Managing Internet Mail • Sybase Performance and Tuning for System 11 • Unix (Beginning) • Unix (Advanced) • Visual Basic 5.0 Fundamentals • Visual C++ Windows Programming

*You do not need an ICN password to use e-mail.

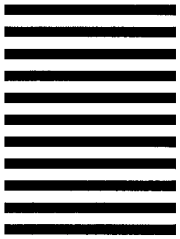
Reader Feedback

Feedback helps us to provide a document that responds to the changing needs of its readership. If you have comments or questions about this publication, please let us hear from you. We have reserved the back of this form for that purpose. We also accept articles for publication that are of interest to our readers. Contact the managing editor for more information. This form is also used for new subscriptions, deletions, or changes. Instructions are on the back. If you prefer to contact us by e-mail, send your comments and/or subscription request to denise@lanl.gov.

Do Not Staple
Fold on This Line First



NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES



BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 88 LOS ALAMOS NM

POSTAGE WILL BE PAID BY THE ADDRESSEE

MAIL STOP B251
ATTN: DENISE SESSIONS, MANAGING EDITOR
CUSTOMER SERVICE GROUP (CIC-6)
LOS ALAMOS NATIONAL LABORATORY
PO BOX 1663
LOS ALAMOS, NM 87545-9916

Do Not Staple, Seal with Tape
Fold Here

cut along dashed line

Feedback

New Subscription, Deletions, and Changes

Bits is published by Los Alamos National Laboratory. If you would like to be added to or deleted from our mailing list, please check the appropppriate line, complete the form below, and mail us the form, or e-mail bitsupdate@lanl.gov

_____ Add my name to the BITS mail list.

_____ Delete my name from the BITS mailing list.

_____ Change my name/address as indicated below.

Name _____ Date _____

Address _____ Mail Stop _____

Group _____ Organization _____

City _____ State _____ Zip _____

Phone _____ Number of copies _____

1999–2000 12-Month Index

To electronically access any of the articles listed on this index go to http://www.lanl.gov/orgs/cic/cic6/bits/index/index_home.html. Note that the page numbers printed on this page correspond to the page numbers in the printed versions of the publication.

Keywords	Title of BITS Article	Date	Page
Networks	*Preserving Bandwidth on Your Windows NT Network*	*F/M 99*	*8*
	News about Networking at Los Alamos: The LANL Master Accounts Domain Is Up and Running!	*Sept. 99*	*10*
	Notes on Setting Up and Configuring an NT Domain	*Sept. 99*	*12*
	New Tool for System Administrators of LANL NT MAD Accounts: Domain Manager Application	*Jan. 00*	*16*
SciSearch	*Research Library Adds Social SciSearch® Database*	*F/M 99*	*5*
Security	*Policy Changes for Foreign Visits and Assignments*	*Jan. 00*	*21*
	Publication Release and Accountability at the Laboratory	*Jan. 00*	*23*
Virtual Private Network	*Virtual Private Network*	*J/J 99*	*15*
Windows NT	*Preserving Bandwith on Your Windows NT Network*	*F/M 99*	*8*
	News about Networking at Los Alamos: The LANL Master Accounts Domain Is Up and Running!	*Sept. 99*	*10*
	Notes on Setting Up and Configuring an NT Domain	*Sept. 99*	*12*
	New Tool for System Administrators of LANL NT MAD Accounts: Domain Manager Application	*Jan. 00*	*16*
World Wide Web	*Lab Web Publishers Invited to IntraLab99*	*F/M 99*	*7*
(WWW or Web)	*Extensible HTML*	*J/J 99*	*1*
	Scalable Vector Graphics—Web Graphics with Original-Quality Artwork	*Nov. 99*	*2*
Writing Resources	*Telling Your Story in BITS: Sharing Your Expertise and Energy*	*Aug. 99*	*31*
Year 2000 (Y2K)	*The Year 2000 Bug Is Hiding*	*F/M 99*	*1*
	Year 2000: Are You Still Ready?	*Nov. 99*	*14*

Note page . . .

Customer Support Center . . . . (505) 665-4444, ext. 851, or cichelp@lanl.gov

Because of a wide variety of CIC computing services, numerous facilities are available to address your questions. If you are uncertain whom to call, you can always call the Customer Support Center (CSC). CSC consultants are trained to either answer your question or locate someone who can. To reach the appropriate consultant, dial 665-4444 and make your selection from the following choices:

Option 1: New user topics including e-mail, passwords, registration, and World Wide Web
Option 2: Enterprise Information Applications such as Travel, Time and Effort, and Purchase cards
Option 3: Scientific computing, storage systems, and networking
Option 4: Classroom instruction and training
Option 5: Desktop Consulting for PC and Macintosh software and network configurations.

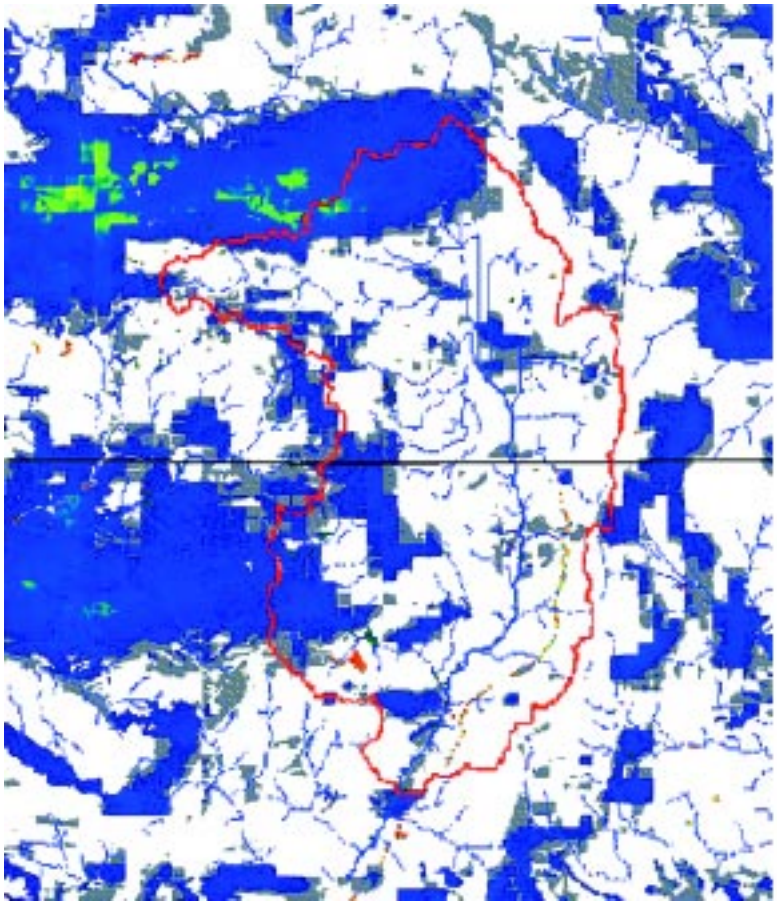
Consulting Via E-Mail

Customer Support ..... cichelp@lanl.gov
Scientific and engineering computing ..... consult@lanl.gov
Administrative and business computing .... eiaconsult@lanl.gov
Passwords and registration ..... validate@lanl.gov
Macintosh computing ..... Mac-help@lanl.gov
PC computing ..... PC-help@lanl.gov
UNIX computing ..... UNIX-help@lanl.gov

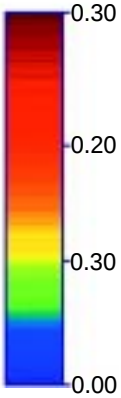
Other Useful Numbers

Advanced Computing Laboratory 665-4530
Central Computing Facility 667-4584
Network Operations Center ..... noc@lanl.gov or 667-7423
Telephone Services Center 667-3400

Los Alamos
NATIONAL LABORATORY



Rio Grande Basin Soil Moisture Day 38



This figure shows the soil moisture at a 100-m resolution for the Rio Grande Basin (outlined in red) from the headwaters in the San Juan Mountains to north of Albuquerque, NM, using the Los Alamos coupled river basin simulator. For more information, send e-mail to Larry Winter, winter@lanl.gov, Computer Research & Applications (CIC-3), or Everett Springer, everetts@lanl.gov, Environmental Science (EES-15). For a Web link about coupled modeling, see http://www.ees15.lanl.gov/capability02a.htm .

Los Alamos
NATIONAL LABORATORY

Mailstop B251 Los Alamos, New Mexico 87545
A U.S. Department of Energy Laboratory

BITS is published bimonthly to highlight recent computing and communications activities within the Laboratory. We welcome your suggestions and contributions.

BITS may be accessed electronically at this URL:
http://www.lanl.gov/cic/publications.html

LALP-00-2 (2-3/00)

Nonprofit organization
US Postage
PAID
Los Alamos, NM
Permit No. 532